Annals of Mathematics
Vol. 35, No. 4, October, 1934

ON THE DEVIATION OF GEODESICS AND NULL-GEODESICS, PARTICULARLY IN RELATION TO THE PROPERTIES OF SPACES OF CONSTANT CURVATURE AND INDEFINITE LINE-ELEMENT

By J. L. Synge

(Received June 12, 1933)

CONTENTS

1. **Introduction.** The object of this paper is to draw attention to the use of the equation of geodesic deviation[1] for the geometrical interpretation of Riemannian curvature and for the exploration of the properties of Riemannian spaces in the large. The fundamental form is not restricted to be positive-definite. The general equation of deviation (2.9) is here deduced in a very simple manner by a method which is valid whether the curves in question are ordinary geodesics or null-geodesics. In (3.11) a simple geometrical interpretation of the Riemannian curvature associated with a 2-element is given in terms of the deviation of geodesics emanating in the 2-element. When the space is of constant curvature, the equation of deviation may be integrated: this is done for ordinary geodesics in (4.6) and (4.12) and for null-geodesics in (4.17) and (4.23). The case where the two geodesics emanate from a point is of particular interest. The paper ends with a brief discussion of the properties in the large of space-time of constant curvature, in so far as these follow from the equation of geodesic deviation.

2. **The equation of deviation.** Let there be a singly infinite family of curves, which are either all ordinary geodesics or all null-geodesics, in a Riemannian space of N dimensions with a fundamental form $a_{ij}\, dx^i\, dx^j$, not necessarily definite. These curves define a V_2. Let v be a parameter which is constant along each curve, and u a parameter which varies along each curve, being equal to the arc length s measured from some curve drawn across the family if the family consists of ordinary geodesics, and being one of a special system of privileged parameters if the family consists of null-geodesics. These parameters are those for which the differential equations of the null-geodesics

[1] T. Levi-Civita, *Math. Ann.* 97 (1926), 315; J. L. Synge, *Phil. Trans. Roy. Soc.* (A) 226 (1926), 102.

take the form (2.3). If T^i is any contravariant vector field defined over V_2, we put

$$\text{(2.1)} \qquad \frac{\delta T^i}{\delta u} = \frac{\partial T^i}{\partial u} + \left\{ {i \atop jk} \right\} T^j \frac{\partial x^k}{\partial u}, \qquad \frac{\delta T^i}{\delta v} = \frac{\partial T^i}{\partial v} + \left\{ {i \atop jk} \right\} T^j \frac{\partial x^k}{\partial v}.$$

Then, as is well known,

$$\text{(2.2)} \qquad \frac{\delta^2 T^i}{\delta u \delta v} - \frac{\delta^2 T^i}{\delta v \delta u} = R^{\,i}_{\cdot\, jkl} T^j \frac{\partial x^k}{\partial u} \frac{\partial x^l}{\partial v},$$

where $R^{\,i}_{\cdot\, jkl}$ is the curvature tensor. From the geodesic character of the family, we have throughout V_2

$$\text{(2.3)} \qquad \frac{\delta}{\delta u} \frac{\partial x^i}{\partial u} = 0,$$

with the first integral

$$\text{(2.4a)} \qquad a_{ij} \frac{\partial x^i}{\partial u} \frac{\partial x^j}{\partial u} = \pm 1,$$

if the family consists of ordinary geodesics, and the first integral

$$\text{(2.4b)} \qquad a_{ij} \frac{\partial x^i}{\partial u} \frac{\partial x^j}{\partial u} = 0,$$

if the family consists of null-geodesics.

Let η^i denote the infinitesimal displacement from a point $P(u, v)$ on a curve C of the family to the point $P'(u, v + dv)$ of a neighbouring curve C' of the family: then

$$\text{(2.5)} \qquad \eta^i = \frac{\partial x^i}{\partial v} dv,$$

and hence

$$\text{(2.6)} \qquad \frac{\delta \eta^i}{\delta u} = \left(\frac{\delta}{\delta u} \frac{\partial x^i}{\partial v} \right) dv = \left(\frac{\delta}{\delta v} \frac{\partial x^i}{\partial u} \right) dv,$$

and, by (2.2),

$$\text{(2.7)} \qquad \begin{aligned} \frac{\delta^2 \eta^i}{\delta u^2} &= \left(\frac{\delta^2}{\delta u \delta v} \frac{\partial x^i}{\partial u} \right) dv \\ &= \left(\frac{\delta^2}{\delta v \delta u} \frac{\partial x^i}{\partial u} - R^{\,i}_{\cdot\, jkl} \frac{\partial x^j}{\partial u} \frac{\partial x^k}{\partial v} \frac{\partial x^l}{\partial u} \right) dv. \end{aligned}$$

But

$$\text{(2.8)} \qquad \frac{\delta^2}{\delta v \delta u} \frac{\partial x^i}{\partial u} = \frac{\delta}{\delta v} \left(\frac{\delta}{\delta u} \frac{\partial x^i}{\partial u} \right) = 0$$

by (2.3), and hence we have the equation of deviation in the form

$$\frac{\delta^2 \eta^i}{\delta u^2} + R^{\,i}_{\cdot\, j k l} \frac{\partial x^j}{\partial u} \eta^k \frac{\partial x^l}{\partial u} = 0 . \tag{2.9}$$

Operating on (2.4a) or (2.4b) with $\delta/\delta v$, we have

$$a_{ij} \left(\frac{\delta}{\partial v} \frac{\partial x^i}{\partial u} \right) \frac{\partial x^j}{\partial u} = 0 ; \tag{2.10}$$

but we have

$$\frac{\delta}{\delta v} \frac{\partial x^i}{\partial u} = \frac{\delta}{\delta u} \frac{\partial x^i}{\partial v} , \tag{2.11}$$

and hence (2.10) may be written

$$\frac{d}{du} \left(a_{ij}\, \eta^i \frac{\partial x^j}{\partial u} \right) = 0 , \tag{2.12}$$

on account of (2.3). Thus, for both ordinary geodesics and null-geodesics,

$$\eta_j \frac{\partial x^j}{\partial u} = \text{const.} \tag{2.13}$$

This constant may be made zero in the case of ordinary geodesics by choosing for curve $u = 0$ an orthogonal trajectory of the family.

3. **Interpretation of Riemannian curvature in terms of geodesic deviation.** Let us consider two adjacent ordinary geodesics C, C' emanating from a point O from which s is measured on both curves. Let us define a unit vector μ^i codirectional with the displacement η^i by means of

$$\mu^i = \eta^i/\eta , \qquad \eta^2 = \epsilon(\mu)\eta_i\eta^i , \qquad \eta > 0 , \tag{3.1}$$

where $\epsilon(\mu)$ is the indicator of η^i, chosen equal to $+1$ or -1 to make

$$\epsilon(\mu)\eta_i\eta^i > 0 .$$

We shall assume that η^i is not a null-vector.

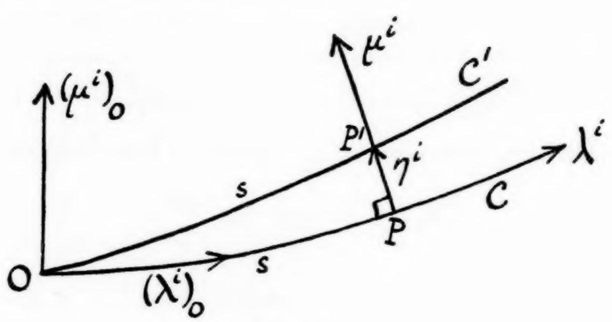

Fig. 1

Let us put $u = s$ in (2.9) and write $\lambda^i = \partial x^i/\partial u$, the unit tangent vector to the geodesic. This gives to the equation of deviation the form

$$\frac{\delta^2 \eta^i}{\delta s^2} + R^{\,i}_{\cdot\,jkl}\,\lambda^j \eta^k \lambda^l = 0\,. \tag{3.2}$$

Let us then substitute for η^i from (3.1): this gives

$$\frac{d^2\eta}{ds^2}\,\mu^i + 2\,\frac{d\eta}{ds}\,\frac{\delta \mu^i}{\delta s} + \eta\,\frac{\delta^2 \mu^i}{\delta s^2} + \eta R^{\,i}_{\cdot\,jkl}\,\lambda^j \mu^k \lambda^l = 0\,. \tag{3.3}$$

Now

$$\epsilon(\mu)\mu_i\mu^i = 1\,, \qquad \mu_i\,\frac{\delta \mu^i}{\delta s} = 0\,, \qquad \mu_i\,\frac{\delta^2 \mu^i}{\delta s^2} + \frac{\delta \mu_i}{\delta s}\,\frac{\delta \mu^i}{\delta s} = 0\,, \tag{3.4}$$

and

$$\lim_{s\to 0} \eta = 0\,, \qquad \lim_{s\to 0} \frac{d\eta}{ds} \neq 0\,. \tag{3.5}$$

Hence, multiplying (3.3) by μ_i and proceeding to the limit $s \to 0$, we obtain

$$\lim_{s\to 0} \frac{d^2\eta}{ds^2} = 0\,. \tag{3.6}$$

Then taking the limit $s \to 0$ in (3.3) we obtain

$$\lim_{s\to 0} \frac{\delta \mu^i}{\delta s} = 0\,, \tag{3.7}$$

showing that μ^i is propagated parallelly at O. Hence, from the last of (3.4),

$$\lim_{s\to 0} \mu_i\,\frac{\delta^2 \mu^i}{\delta s^2} = 0\,, \tag{3.8}$$

and so, if we multiply (3.3) by μ_i/η and proceed to the limit $s \to 0$, we obtain

$$\epsilon(\mu)\,\lim_{s\to 0}\,\eta\,\frac{d^2\eta}{ds^2} + \lim_{s\to 0} R_{ijkl}\,\mu^i\lambda^j\mu^k\lambda^l = 0\,. \tag{3.9}$$

Now if λ^i, μ^i are perpendicular unit vectors at a point, the Riemannian curvature associated with the 2-element containing them is defined to be[2]

$$K(\lambda, \mu) = \epsilon(\lambda)\epsilon(\mu)\,R_{ijkl}\,\lambda^i\mu^j\lambda^k\mu^l\,, \tag{3.10}$$

where $\epsilon(\lambda)$, $\epsilon(\mu)$ are the indicators of the vectors. Thus (3.9) gives

$$K(\lambda, \mu) = -\epsilon(\lambda)\,\lim_{s\to 0} \frac{1}{\eta}\,\frac{d^2\eta}{ds^2}\,, \tag{3.11}$$

[2] This is equivalent to the definition more commonly employed: cf. Eisenhart, Riemannian Geometry, p. 81.

where $K(\lambda, \mu)$ is the Riemannian curvature of the two-dimensional element in which the geodesics emanate. This gives a geometrical interpretation of Riemannian curvature. If η is expanded in a series

$$\eta = a_1 s + \tfrac{1}{2}a_2 s^2 + \tfrac{1}{6}a_3 s^3 + \cdots , \tag{3.12}$$

we have by (3.6) $a_2 = 0$, and by (3.11)

$$K(\lambda, \mu) = -\epsilon(\lambda) a_3/a_1 . \tag{3.13}$$

4. **Deviation in a space of constant curvature.** In a space of constant curvature K we have

$$R_{ijkl} = K(a_{ik}\, a_{jl} - a_{il}\, a_{jk}) . \tag{4.1}$$

Substitution in (2.9) gives

$$\frac{\delta^2 \eta^i}{\delta u^2} + K\left(\eta^i\, a_{jk} \frac{\partial x^j}{\partial u} \frac{\partial x^k}{\partial u} - \frac{\partial x^i}{\partial u} a_{jk}\, \eta^j \frac{\partial x^k}{\partial u}\right) = 0 . \tag{4.2}$$

For the case of a family of ordinary geodesics (with the orthogonal correspondence assigned at the end of §2) this becomes (putting s for u)

$$\frac{\delta^2 \eta^i}{\delta s^2} + \epsilon(\lambda) K \eta^i = 0 , \tag{4.3}$$

where $\epsilon(\lambda)$ is the indicator of the geodesics, chosen equal to $+1$ or -1 to make

$$\epsilon(\lambda) a_{ij} \frac{\partial x^i}{\partial u} \frac{\partial x^j}{\partial u} > 0 .$$

For a family of null-geodesics, on the other hand, we have

$$\frac{\delta^2 \eta^i}{\delta u^2} - K\left(\eta_k \frac{\partial x^k}{\partial u}\right) \frac{\partial x^i}{\partial u} = 0 . \tag{4.4}$$

To integrate (4.3), we let ν^i be any vector propagated parallelly along C. Then, multiplying (4.3) by ν_i, we have

$$\frac{d^2}{ds^2} (\eta^i \nu_i) + \epsilon(\lambda) K \eta^i \nu_i = 0 , \tag{4.5}$$

and hence

$$\left\{\begin{aligned} \eta^i \nu_i &= A \sin s[\epsilon(\lambda)K]^{\frac{1}{2}} + B \cos s[\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K > 0 , \\ \eta^i \nu_i &= As + B \text{ if } K = 0 , \\ \eta^i \nu_i &= A \sinh s[-\epsilon(\lambda)K]^{\frac{1}{2}} + B \cosh s[-\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K < 0, \end{aligned}\right. \tag{4.6}$$

where A, B are constants. This shows that, if $\epsilon(\lambda)K > 0$, η^i either vanishes in all its components or becomes perpendicular to ν^i periodically, the period in s being $\pi[\epsilon(\lambda)K]^{-\frac{1}{2}}$.

Another method of integrating (4.3) is to multiply by η_i, and by $\delta\eta_i/\delta s$, obtaining

$$\eta_i \frac{\delta^2\eta^i}{\delta s^2} + \epsilon(\lambda)K\eta_i\eta^i = 0 , \tag{4.7}$$

$$\frac{\delta\eta_i}{\delta s}\frac{\delta^2\eta^i}{\delta s^2} + \epsilon(\lambda)K\frac{\delta\eta_i}{\delta s}\eta^i = 0 . \tag{4.8}$$

Equation (4.8) gives on integration

$$\frac{\delta\eta_i}{\delta s}\frac{\delta\eta^i}{\delta s} + \epsilon(\lambda)K\eta_i\eta^i = C , \tag{4.9}$$

where C is a constant. But

$$\frac{d^2}{ds^2}(\eta_i\eta^i) = 2\eta_i\frac{\delta^2\eta^i}{\delta s^2} + 2\frac{\delta\eta_i}{\delta s}\frac{\delta\eta^i}{\delta s} , \tag{4.10}$$

or, on substitution on the right from (4.7) and (4.9),

$$\frac{d^2}{ds^2}(\eta_i\eta^i) = 2C - 4\epsilon(\lambda)K\eta_i\eta^i ; \tag{4.11}$$

hence

$$\begin{cases} \eta^2 = C' + A'\sin 2s[\epsilon(\lambda)K]^{\frac{1}{2}} + B'\cos 2s[\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K > 0 , \\ \eta^2 = \epsilon(\mu)\, Cs^2 + A's + B' \text{ if } K = 0 , \\ \eta^2 = C' + A'\sinh 2s[-\epsilon(\lambda)K]^{\frac{1}{2}} + B'\cosh 2s[-\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K < 0 , \end{cases} \tag{4.12}$$

where $C' = \frac{1}{2}\epsilon(\lambda)\epsilon(\mu)C/K$, and A', B' are constants. *Equations* (4.6) *and* (4.12) *express the deviation of ordinary geodesics in a space of constant curvature.*

Turning now to the equation (4.4) for the deviation of null-geodesics, we have by (2.13)

$$\eta_j\frac{\partial x^j}{\partial u} = C , \tag{4.13}$$

a constant. Hence we have

$$\frac{\delta^2\eta^i}{\delta u^2} - KC\frac{\partial x^i}{\partial u} = 0 . \tag{4.14}$$

We may proceed in two ways to integrate this. Let ν^i be any vector propagated parallelly along the null-geodesic. Then

$$\nu_i\frac{\partial x^i}{\partial u} = D , \tag{4.15}$$

a constant, and we obtain from (4.14)

$$\frac{d^2}{du^2}(\nu_i\eta^i) = KC\nu_i\frac{\partial x^i}{\partial u} = KCD , \tag{4.16}$$

which gives

$$\nu_i \eta^i = \tfrac{1}{2} u^2 KCD + Au + B\,, \tag{4.17}$$

where A, B are constants. On the other hand, we may proceed as follows: if we differentiate (4.13), we obtain, by (2.3),

$$\frac{\delta \eta_j}{\delta u} \frac{\partial x^j}{\partial u} = 0\,, \tag{4.18}$$

and hence, multiplying (4.14) by η_i and by $\delta\eta_i/\delta u$, we have, by (4.13),

$$\eta_i \frac{\delta^2 \eta^i}{\delta u^2} = KC^2\,, \tag{4.19}$$

$$\frac{\delta \eta_i}{\delta u} \frac{\delta^2 \eta^i}{\delta u^2} = 0\,. \tag{4.20}$$

The latter gives on integration

$$\frac{\delta \eta_i}{\delta u} \frac{\delta \eta^i}{\delta u} = C''\,, \tag{4.21}$$

a constant. If we now write down the identity (4.10), with u instead of s, and substitute on the right from (4.19) and (4.21), we obtain

$$\frac{d^2}{du^2} (\eta_i \eta^i) = 2KC^2 + 2C''\,, \tag{4.22}$$

and hence

$$\eta^2 = C'u^2 + A'u + B'\,, \tag{4.23}$$

where A', B' are constants and $C' = \epsilon(\mu)(KC^2 + C'')$, $\epsilon(\mu)$ being the indicator of η^i. *Equations* (4.17) *and* (4.23) *express the deviation of null-geodesics in a space of constant curvature.*

Let us apply (4.6) to the case where the ordinary geodesics emanate from a point, at which $s = 0$. Then $B = 0$, and we have

$$\begin{cases} \eta^i \nu_i = A \sin s\,[\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K > 0\,, \\ \eta^i \nu_i = As \text{ if } K = 0\,, \\ \eta^i \nu_i = A \sinh s\,[-\epsilon(\lambda)K]^{\frac{1}{2}} \text{ if } \epsilon(\lambda)K < 0\,. \end{cases} \tag{4.24}$$

Remembering that the components ν_i may be arbitrarily assigned at any one point, we see that η^i all vanish, not only for $s = 0$, but also for $s = \pi[\epsilon(\lambda)K]^{-\frac{1}{2}}$ in the first case. Hence *in a space of constant curvature K adjacent geodesics emanating from a point in a direction whose indicator $\epsilon(\lambda)$ has the same sign as K meet again after a distance $\pi \mid K \mid^{-\frac{1}{2}}$, but if they emanate in a direction whose indicator has the opposite sign to K, they diverge exponentially.*

Passing continuously from geodesic to geodesic, we see that *in a space of constant curvature K all geodesics emanating from a point O in directions whose indicators have the same sign as K, and which form a connected vector-space, meet again at a common distance $\pi \mid K \mid^{-\frac{1}{2}}$ from O.*

Let us apply (4.17) to the case where the null-geodesics emanate from a point O where $u = 0$. Since $\eta^i = 0$ at O, we have, by (4.13), $C = 0$, and by (4.17), $B = 0$. Hence (4.17) becomes

$$\nu_i \eta^i = Au , \tag{4.25}$$

where the components ν_i may be arbitrarily assigned at any one point of the null-geodesic. We may also deduce from (4.23)

$$\eta^2 = C' u^2 . \tag{4.26}$$

Hence adjacent null-geodesics emanating from a point in a space of constant curvature diverge linearly in terms of the special parameter u.

5. Properties of space-time of constant curvature in the large. Let us consider a space of 4-dimensions (space-time) whose fundamental form has the signature

$$+ + + - . \tag{5.1}$$

The elementary null-cone at a point O divides the vector field at O into three regions:

(i) the future,
(ii) the present,
(iii) the past,

the present being "exterior" to the null cone and having an indicator $+1$ for vectors drawn from O into it (space-like directions), and the future and past being "interior" to the null cone and having an indicator -1 for vectors drawn from O into them (time-like directions). The future, present, and past are individually connected vector-spaces, but the future and the past are not connected.

Hence if space-time (with the signature (5.1)) *is of constant positive curvature K, all geodesics drawn into the present meet again at a common distance $\pi K^{-\frac{1}{2}}$, but those drawn into the future or past diverge exponentially.*

On the other hand *if space-time (with the signature* (5.1)) *is of constant negative curvature K, geodesics drawn into the present diverge exponentially, but all geodesics drawn into the future meet again at a common distance $\pi(-K)^{-\frac{1}{2}}$, and all those drawn into the past meet again at a common distance $\pi(-K)^{-\frac{1}{2}}$.*

We have seen that if K is positive all geodesics emanating from O into the present converge on a point O'. It is easily seen that these geodesics (which fill the present at O) also fill the present at O'. For, if they do not, let P be a point in the present of O' near O', not on a geodesic of the family, but close to a member C of the family. Let $(\eta^i)_P$ be the infinitesimal displacement NP

from C to P, $(\eta^i)_P$ being perpendicular to C. But we can find solutions η^i of (4.3) taking these assigned values $(\eta^i)_P$ at N, and vanishing at O and O'. This will define a geodesic joining O and O', not in the family, contrary to hypothesis. Hence the geodesics in question fill the present at O'. Each geodesic from O, arriving at O', finds facing it another geodesic of the family from O; travelling back along this latter geodesic we arrive at O again. Hence *every space-like geodesic in a space-time of constant positive curvature and signature* (5.1) *is a closed curve, whose length is* $2\pi K^{-\frac{1}{2}}$ *if* O' *is not the same point at* O, *and* $\pi K^{-\frac{1}{2}}$ *if* O' *is the same point as* O. (The space is *polar* if O' coincides with O, and *antipodal* if O' does not coincide with O. It does not follow in either case that when we pass through O for the second time the direction is the same as in the first passage through O.)

If we make a postulate of symmetry, to the effect that it is impossible intrinsically to distinguish between two space-like directions at a point, or between two time-like directions at a point, or between two null-directions at a point, it is easily seen that for positive K no space-like geodesic from O can pass through O again until it has described a distance either $\pi K^{-\frac{1}{2}}$ or $2\pi K^{-\frac{1}{2}}$, and no time-like geodesic from O can ever pass through O again. Nor can a null-geodesic from O pass through O again.

Let us now briefly consider the case where K is negative and where, consequently, all time-like geodesics from O converge on O', at a distance $\pi \mid K \mid^{-\frac{1}{2}}$. Obviously they fill the past at O', as they filled the future at O. But when they pass on through O', they cannot pass back along their own system, for they pass into the future at O'. Hence we cannot deduce that the time-like geodesics for negative K are necessarily closed, as we did for space-like geodesics for positive K.

UNIVERSITY OF TORONTO.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

ON GENERALIZED RIEMANN MATRICES

By Hermann Weyl

(Received May 10, 1934)

In the following I intend to give a simpler and generalized formulation of the problem of complex multiplication of Riemann matrices, recently treated with such conclusive success by A. A. Albert.[1] All known propositions remain untouched by this generalization, which in my opinion is required by the nature of the subject.

§1. Foundations: Transfer of the function theoretical problem into the algebraic one

On a Riemann surface of connectivity $n = 2p$ (genus p), we take a basis α of the closed curves. Every closed curve ξ is homologous to a linear combination of the n basic curves, $\xi \sim \sum_\alpha x_\alpha \cdot \alpha$, by means of integers x_α. The "characteristic" $[\xi\eta]$ of any two closed curves ξ and $\eta \sim \sum_\alpha y_\alpha \cdot \alpha$ which gives the number of times ξ crosses η in summa in the positive sense is a non-degenerate skew-symmetric bilinear form

$$[\xi\eta] = \sum_{\alpha,\beta} c_{\alpha\beta} x_\alpha y_\beta\,; \qquad c_{\alpha\beta} = [\alpha\beta]$$

with integer coefficients. Transition to a new basis α of curves is performed by a unimodular integral transformation U. The construction of the integrals of the first kind through Dirichlet's principle naturally leads to associating a differential of the first kind dw_α with every closed curve α such that, for every closed curve β the real part $\Re \int_\beta dw_\alpha = c_{\alpha\beta} = [\alpha\beta]$ equals the characteristic of the two curves α and β.[2] Homologous curves α are associated with the same dw_α, addition of curves α leads to addition of the corresponding differentials dw_α. In this manner the basis of curves α gives rise to a "real basis" dw_α for the differentials of the first kind, consisting of n terms; every differential of the first kind dw can be uniquely expressed as a linear combination

$$dw = \sum_\alpha x_\alpha \cdot dw_\alpha \tag{1}$$

[1] Rend. Circ. Mat. di Palermo **55** (1931), p. 57; Trans. Amer. Math. Soc. **33** (1931), p. 219; Annals of Math. **35** (1934), p. 1, to be continued. All work prior to 1928, in particular by Scorza, Rosati and Lefschetz, is reported in the latter's Report of the Committee on Rational Transformations, Bulletin of the National Research Council, **63** (1928), pp. 310–392.

[2] Weyl, Idee der Riemannschen Fläche, 2[te] Auflage, Leipzig, 1923, p. 98 and pp. 172–174. A general topological proof of the fact that the characteristic form is non-degenerate and even unimodular, which remains valid for higher dimensions and does not pass through the explicit construction of a canonical basis: Weyl, Revista Mat. Hispano-Americana, 1923, Theorem 10.

with constant real coefficients x_α. The basis dw_α transforms cogrediently with the basis α of the curves themselves. Whereas the real parts of the periods

$$\Re \int_\beta dw_\alpha = c_{\alpha\beta}$$

form an integral, non-singular, skew-symmetric matrix C, the imaginary parts

$$\Im \int_\beta dw_\alpha = s_{\alpha\beta}$$

are symmetric and the coefficients of a positive definite quadratic form. The definite character of the quadratic form $S = \Sigma\, s_{\alpha\beta}\, x_\alpha\, x_\beta$ may be described as the property of being non-singular in every real partial space of the total n-dimensional vector space. This is meant if we say that S is *totally regular in the field of real numbers.*[3]

When one operates with a real basis it is natural to ask how the differentials multiplied by i are expressed by the differentials themselves in a real manner:

$$idw_\alpha = \sum\nolimits_\gamma r_{\gamma\alpha}\, dw_\gamma \qquad (r_{\gamma\alpha} \text{ real constants}). \tag{2}$$

By integrating over β and taking the real part, one obtains the equations

$$-s_{\alpha\beta} = \sum\nolimits_\gamma r_{\gamma\alpha}\, c_{\gamma\beta} \qquad \text{or} \qquad s_{\beta\alpha} = \sum\nolimits_\gamma c_{\beta\gamma}\, r_{\gamma\alpha}.$$

Consequently the relation

$$S = C \cdot R, \qquad R = C^{-1}S \tag{3}$$

holds for the matrices

$$C = \| c_{\alpha\beta} \|, \qquad S = \| s_{\alpha\beta} \| \qquad \text{and} \qquad R = \| r_{\alpha\beta} \|.$$

The transformation R has, according to its significance (2), the property

$$R^2 = -1. \tag{4}$$

C and S occur in the problem of complex multiplication only in this combination R; and the only assumption concerning R that really matters is that R arises according to (3) from an arbitrary rational non-singular skew-symmetric matrix C and an arbitrary real symmetric totally-regular matrix S. The equation (4) does not play any part and will be discarded. Our generalization in comparison with the formulation used before consists exactly in wiping out this restriction $R^2 = -1$ (compare Appendix, §6).

The question of complex multiplication arises, for instance, when we consider an arbitrary (μ, ν)-valued correspondence $P \rightarrow Q$ on our Riemann surface. P determines the point Q ν-valued: $Q_1, \cdots, Q_\nu$.

$$dw_\alpha(Q_1) + \cdots + dw_\alpha(Q_\nu)$$

is a differential of the first kind with respect to P, let us say

$$= \sum\nolimits_\gamma h_{\alpha\gamma} \cdot dw_\gamma(P).$$

[3] See Weyl loc. cit., p. 116. The form S is the Dirichlet integral of the general differential of the first kind (1) and hence positive.

If P runs over the cycle β, then $Q_1, \cdots, Q_\nu$ together travel over a certain cycle $\sum_\gamma a_{\gamma\beta}\cdot\beta$ (a integers). Hence as $\int_\beta dw_\alpha = c_{\alpha\beta} + is_{\alpha\beta}$, we have in an obvious notation:

$$(C + iS)A = H(C + iS).$$

This splits into the two real equations

$$CA = HC \qquad \text{or} \qquad H = CAC^{-1} \text{ and}$$
$$SA = HS.$$

By substituting H from the first equation, the latter furnishes

$$C^{-1}SA = AC^{-1}S \qquad \text{or} \qquad RA = AR.$$

Let us replace the field of real numbers by any field P and the field of rational numbers by a subfield ρ of P; ρ is considered as the basic domain of rationality. Then we are concerned with the following problem:

Given a matrix R in P arising by equation (3) from a symmetric totally-regular matrix S in P and a skew-symmetric, non-singular matrix C in ρ, the algebra $\mathfrak{A}$ of all matrices A in ρ commuting with R is to be investigated. (In particular we should like to know how a *"Riemann matrix"* R looks, whose *"commutator algebra"* $\mathfrak{A}$ does not consist merely of the multiples of the unit matrix. Hence the problem is to investigate the structure of $\mathfrak{A}$ independently of R and then to find Riemann matrices R corresponding to a given $\mathfrak{A}$ of the ascertained structure.) One may now forget all except this problem.

Only transformations U whose coefficients lie in ρ ("rational transformations") of the coordinate systems are admissible. U carries C, S, R over into

$$U'CU, \qquad U'SU, \qquad U^{-1}RU.$$

§2. Poincaré's and Schur's theorems

POINCARÉ'S THEOREM: *A Riemann matrix R of the reduced form* (5) *can be completely decomposed into its parts R_1 and R_2 by means of an appropriate rational transformation U:*

$$R = \begin{array}{|c|c|}\hline R_1 & 0 \\ \hline Q & R_2 \\ \hline\end{array} \rightarrow \begin{array}{|c|c|}\hline R_1 & 0 \\ \hline 0 & R_2 \\ \hline\end{array} \text{ by } U = \begin{array}{|c|c|}\hline 1 & 0 \\ \hline B & 1 \\ \hline\end{array}. \tag{5}$$

("Rational" always means: lying in ρ.)

One has to prove that Q can be brought into the form

$$Q = BR_1 - R_2B \tag{6}$$

with a rational B. As a hypothesis to start with, we have the equation (3) or

$$\begin{array}{|c|c|}\hline C_{11} & C_{12} \\ \hline C_{21} & C_{22} \\ \hline\end{array} \cdot \begin{array}{|c|c|}\hline R_1 & 0 \\ \hline Q & R_2 \\ \hline\end{array} = \begin{array}{|c|c|}\hline S_{11} & S_{12} \\ \hline S_{21} & S_{22} \\ \hline\end{array}$$

at our disposal. This contains in particular:

$$C_{22} R_2 = S_{22}. \tag{7}$$

S_{22} is non-singular as S is totally regular:

$$\det (S_{22}) \neq 0, \qquad \det (C_{22}) \neq 0.$$

Since C_{22} is skew-symmetric and S_{22} is symmetric, (7) yields by going over to the transposed matrices: $-R_2' C_{22} = S_{22}$,

$$R_2' = -S_{22} C_{22}^{-1} = -C_{22} R_2 C_{22}^{-1}. \tag{8}$$

Furthermore, we have

$$S_{12} = C_{12} R_2, \qquad S_{21} = C_{21} R_1 + C_{22} Q. \tag{9}$$

The symmetry of S and the skew symmetry of C imply

$$S_{12}' = S_{21} \quad \text{and} \quad C_{12}' = -C_{21}.$$

Hence, according to (9):

$$-R_2' C_{21} = C_{21} R_1 + C_{22} Q.$$

In replacing R_2' here by the expression (8) we get, after cancellation of the factor C_{22} in front:

$$R_2 C_{22}^{-1} C_{21} = C_{22}^{-1} C_{21} R_1 + Q.$$

Hence $B = -C_{22}^{-1} C_{21}$ satisfies the desired equation (6).

The proof has not even made use of the fact that S is totally regular in P but only in ρ; that is to say, the quadratic form S is supposed to be non-singular in any partial space spanned by vectors the components of which are numbers in ρ.

R is reducible if it can be rationally transformed into a reduced matrix like (5). $R \sim$ (equivalent) R' means that the matrix R' can be brought into coincidence with R through rational transformation. Following Poincaré's theorem, R may be decomposed into irreducible constituents $R_1, R_2, \cdots$, and it is allowed to assume that the equivalent ones among them are equal; in this manner R breaks up into "blocks" of equal irreducible parts. Some information about the rational commutators A of the splitting Riemann matrix R is provided by

SCHUR'S LEMMA: 1) *R_1 and R_2 being irreducible and inequivalent, zero is the only matrix A in ρ (of the right number of rows and columns) which satisfies the equation*

$$R_1 A = A R_2. \tag{10}$$

2) *Matrices A in ρ which commute with an irreducible R, are either zero or non-singular.*

While A. A. Albert has to use an analogue of Schur's lemma in his treatment, which he proves in a manner similar to Schur's, we are led directly to a particu-

lar case of Schur's lemma as it is known in the theory of representations where the single matrix R is replaced by a whole system, in particular a group of matrices.

The commutators A of R, lying in ρ, split up, in consequence of part 1) of Schur's lemma, into *blocks* corresponding to the blocks of equal irreducible parts of R. For an individual block, however, like

$$\left\| \begin{matrix} R_1 & & \\ & R_2 & \\ & & R_3 \end{matrix} \right\| \qquad (R_1 = R_2 = R_3 = \text{an irreducible } R^{\cdot}) \tag{11}$$

the commutators are of the shape

$$A = \left\| \begin{matrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{matrix} \right\| \tag{12}$$

where the A_{ik} are arbitrary commutators of the irreducible $R^{\cdot}$ (Scorza's theorem).

The fact that $S = CR$ is symmetric finds its expression in the equation

$$CR = R'C' = -R'C. \tag{13}$$

Corresponding to the decomposition of R into irreducible parts like (11), we write every matrix A in the form (12). Since S is totally regular, the parts S_{11}, S_{22}, S_{33} cannot be singular. According to $S = CR$ or $S_{11} = C_{11} R_1, \cdots$, the same holds for C_{11}, C_{22}, C_{33}. The equations

$$C_{11} R_1 = - R_1' C_{11}, \cdots$$

which are contained in (13) then show that $-R_i'$ is rationally equivalent to R_i. Hence, according to part 1) of Schur's lemma and to equation (13), the matrix C splits up into blocks in the manner described before for the commutators A. For an individual block C has the form

$$\left\| \begin{matrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{matrix} \right\|.$$

C stays skew-symmetric and non-singular, and the corresponding $S = CR$ stays symmetric and totally regular if one cancels the lateral terms C_{ik} $(i \neq k)$, i.e. if one now chooses C equal to

$$\left\| \begin{matrix} C_{11} & 0 & 0 \\ 0 & C_{22} & 0 \\ 0 & 0 & C_{33} \end{matrix} \right\|.$$

Following these results we may restrict ourselves for further investigation to an *irreducible* R. The rational commutators A of an irreducible R form according to part 2) of Schur's lemma, a *division algebra* $\mathfrak{A}$.

§3. The involution of the commutator algebra $\mathfrak{A}$.

Any rational commutator A of the irreducible R gives rise by dint of the equation $CA = B$, to a matrix B satisfying the equation

$$BR = -R'B. \tag{14}$$

Conversely, any rational B satisfying this equation is of the form CA, A being a commutator. Transposition of (14) yields this same relation for B'. Hence B' must equal CA^* where the second factor A^* again is a commutator:

$$CA^* = (CA)' = A'C' = -A'C.$$

Thus, after canceling the minus sign, we find that every commutator A is associated with a "dual" one

$$A^* = C^{-1}A'C. \tag{15}$$

The transition $A \rightarrow A^*$ is first an *anti-automorphism*:

$$(A_1 A_2)^* = A_2^* A_1^*,$$

second, an *involution*. For the equation $CA^* = A'C$ changes by transposition into

$$-A^{*\prime}C = -CA \qquad \text{or} \qquad A^{**} = A.$$

We have thus proved the theorem of *Rosati*:[4]

The division algebra $\mathfrak{A}$ permits an anti-automorphic involution $A \rightarrow A^*$.

It may be observed that all previous results remain valid if C is a *symmetric* rather than an anti-symmetric, non-singular matrix in ρ; in this case the number of dimensions n need not be even.

The matrix algebra $\mathfrak{A}$ can be considered as the representation of an abstract division algebra $\mathfrak{a}$ in ρ the fundamental operations of which are: addition and multiplication of quantities of $\mathfrak{a}$, multiplication of a quantity a of $\mathfrak{a}$ by a number of ρ, and involution $a \rightarrow a^*$. The latter has the properties

$$(a + b)^* = a^* + b^*, \qquad (\lambda a)^* = \lambda a^*, \qquad (ab)^* = b^*a^* \quad (\lambda = \text{number in } \rho).$$

We call this an *involutorial division algebra*. $\mathfrak{a}$ has only one irreducible representation $\mathfrak{A}$: $a \rightarrow A = A(a)$ in ρ. The most general representation of $\mathfrak{a}$, and hence in particular the matrix algebra with which we dealt and which may be denoted by $\underline{\mathfrak{A}}$ for the rest of this section, is a multiple of $\mathfrak{A}$; the matrix $\underline{\mathfrak{A}}$ associated with the quantity a in $\underline{\mathfrak{A}}$ splits into t times the matrix A.

[4] Rend. Circ. Mat. di Palermo **53** (1929), p. 79–134.

Hence we proceed as follows: we start with an involutorial division algebra $\mathfrak{a}$ in ρ and its irreducible representation $\mathfrak{A}$. In $\mathfrak{A}$ the dual element a^* may be associated with the matrix A^*. $a^* \to A'$ as well as $\mathfrak{A}$: $a^* \to A^*$ defines a representation of $\mathfrak{a}$; so the former must be equivalent to $\mathfrak{A}$, i.e. there exists a non-singular matrix C_0 in ρ satisfying the equation

$$A(a^*) = C_0^{-1} A'(a) C_0$$

identically with respect to a. The most general matrix C which fulfills the same equation

$$CA^* = A'C \tag{16}$$

identically in a is of the form $C = C_0 L$ where L commutes with all matrices A of $\mathfrak{A}$. The algebra of these L may be denoted by $\mathfrak{L}$, more exactly by $\mathfrak{L}_\rho$ or $\mathfrak{L}_P$ according to whether we suppose that L lies in ρ or P. $\mathfrak{L}_\rho$ is a division algebra on account of the irreducibility of $\mathfrak{A}$, and consequently any C satisfying the equation (16) is either 0 or non-singular.

If one changes the equation (16) for $C = C_0$ to the transposed one, and exchanges A with A^*—which one is justified in doing because of the involutorial character of the mapping $a \to a^*$—one sees that C_0' satisfies the same equation. For this reason one can choose C_0 *either symmetric or skew-symmetric*. For if C_0 is not symmetric, one forms $C_0 - C_0' = C^0$; this is a solution of (16), $\neq 0$ and hence, according to our above remark, non-singular, so we can use C^0 instead of C_0.

Let C_0 be symmetric or skew-symmetric from now on. The fact that the equation (16) is satisfied by C' if by C, shows that the matrix $L^* = C_0^{-1} L' C_0$ always lies in $\mathfrak{L}$ if L does. L may be called *even* or *odd* according as $L^* = \pm L$. Any L is the sum of an even and an odd L. If we start with the irreducible matrix algebra $\mathfrak{A}$, the corresponding matrices C and S have to be of the form

$$C = C_0 L, \qquad S = C_0 M,$$

where L and M are even or odd matrices in $\mathfrak{L}_\rho$ and $\mathfrak{L}_P$ respectively,—even or odd according to whether or not C and S are to be of the same parity as C_0.

$$R = C^{-1} S = L^{-1} M$$

is a matrix in $\mathfrak{L}_P$ which can be obtained in this manner from an even or odd matrix L in $\mathfrak{L}_\rho$ and an even or odd matrix M in $\mathfrak{L}_P$.

If $\underline{\mathfrak{A}}$ equals t times the irreducible representation $\mathfrak{A}$, we choose as our $\mathbf{C}_0$ the matrix that decomposes into t times C_0. The most general matrix $\mathbf{L}$ of $\underline{\mathfrak{L}}$ is of shape

$$\begin{Vmatrix} L_{11} & \cdots & L_{1t} \\ \cdots & \cdots & \cdots \\ L_{t1} & \cdots & L_{tt} \end{Vmatrix}$$

where the L_{ik} are arbitrary matrices in $\mathfrak{L}$. $\mathbf{L}^* = \mathbf{C}_0^{-1}\, \mathbf{L}'\, \mathbf{C}_0$. $\mathbf{L}'$ will be even: $\mathbf{L} = \mathbf{L}^*$, if the L_{ii} along the main diagonal are even, and if equations like $L_{21} = L_{12}^*$ hold for the lateral terms. $\mathbf{R} = \mathbf{L}^{-1}\,\mathbf{M}$; $\mathbf{L}$ in $\mathfrak{L}_\rho$, $\mathbf{M}$ in $\mathfrak{L}_P$, $\mathbf{L}$ and $\mathbf{M}$ even or odd.

The commutator algebra of the $\mathbf{R}$, to which we are led in this way, comprises the given $\underline{\mathfrak{A}}$ without being necessarily identical with it. The general solution $\mathbf{R}$, however, will depend on certain parameters, and one might expect that the commutator algebra will not embrace more than $\underline{\mathfrak{A}}$ if these parameters avoid certain special conditions.

The question as to whether the rational non-singular C_0, which by means of (15) effects the transition $A \to A^*$ is to be chosen symmetrically or skew-symmetrically, can be decided by more refined tools only,—at least in the case of even dimensionality (compare section 5). For odd dimensionality, of course, only the case of a symmetric C_0 can occur. For a two-term *reducible* representation $\underline{\mathfrak{A}}$ ($t = 2$) both possibilities may be arrived at: for one may put

$$\mathbf{C}_0 = \begin{Vmatrix} 0 & C_{12} \\ C_{21} & 0 \end{Vmatrix}$$

and take as C_{12} a rational non-singular solution of (16) and $C_{21} = \pm\, C_{12}'$ according as one wishes a symmetric or a skew-symmetric $\mathbf{C}_0$. The same remark holds in general for even t.

The result of these considerations is: that *our problem can be reduced essentially to the construction of all involutorial division algebras in ρ.*

For more detailed analysis one will have recourse to the *splitting fields* of the division algebra $\mathfrak{a}$ and the corresponding factor systems according to I. Schur and R. Brauer.[5] This method even before it yields the algebra $\mathfrak{a}$ and its representation $\mathfrak{A}$, leads to the algebra $\mathfrak{L}$ of the matrices commutable with $\mathfrak{A}$ in which the Riemann matrix R lies.

§4. Adjunction of the centrum

From now on we discard the use of the bold-face symbols: unless otherwise stated, $\mathfrak{A}$ denotes either the irreducible representation of $\mathfrak{a}$ or a multiple of it. Following Schur (loc. cit.) one may proceed as follows. Let A be a matrix of the rational commutator algebra $\mathfrak{A}$. The characteristic polynomial $|\, \lambda 1 - A\, |$ of A (1 = unit matrix, λ the variable) shall be decomposed into its irreducible factors in ρ: $\prod \varphi(\lambda)$. For an individual factor $\varphi(\lambda)$ of degree h, one has $|\, \varphi(A)\, | = 0$ and hence $\varphi(A) = 0$. Let us start in the vector space of the transformations A with a vector $e \neq 0$ and then form the series $e, Ae, A^2e, \cdots$. They span an h-dimensional partial space invariant with respect to A, in which the transformation A has the characteristic polynomial $\varphi(\lambda)$ with its roots

[5] Schur, Trans. Amer. Math. Soc. (2) **15** (1909), p. 159. Brauer, Math. Zschr. **28** (1928), p. 67.

$\alpha_1, \cdots, \alpha_h$ all different from each other. By repeating the construction a second time for a vector e' not contained in the partial space thus found, a third time, $\cdots$, finally a ν^{th} time, one breaks up the whole vector space in a number of partial spaces of the described kind. Hence the characteristic polynomial in the total vector space equals $\{\varphi(\lambda)\}^\nu$, and in an appropriate co-ordinate system A becomes a diagonal matrix along the main diagonal of which appears t times the number α_1, then t times the number α_2, and so on. As one readily sees from the equation $RA = AR$, the matrix R splits up in this (non-rational!) co-ordinate system into h matrices of order ν:

$$R = \left\| \begin{matrix} R_1 & & \\ & \ddots & \\ & & R_h \end{matrix} \right\|. \tag{17}$$

Let P be again, and from now on, the field of all real numbers. By performing the transformation into the co-ordinate system just introduced, possibly not real, the matrices C and S shall be treated as the coefficient matrices of bilinear forms of the variables $\bar{x}_\alpha$ and y_α:

$$\sum_{\alpha\beta} c_{\alpha\beta}\, \bar{x}_\alpha\, y_\beta, \qquad \sum_{\alpha\beta} s_{\alpha\beta}\, \bar{x}_\alpha\, y_\beta,$$
$$C \to \bar{U}'CU, \qquad S \to \bar{U}'SU.$$

$\sum s_{\alpha\beta}\, \bar{x}_\alpha\, x_\beta$ then remains a definite Hermitian form; the conditions of symmetry read $\bar{C}' = -C$, $\bar{S}' = S$. In agreement with (17) we write, with reference to the new co-ordinate system,

$$C = \left\| \begin{matrix} C_{11} & \cdots & C_{1h} \\ \cdots & \cdots & \cdots \\ C_{h1} & \cdots & C_{hh} \end{matrix} \right\|, \qquad S = \left\| \begin{matrix} S_{11} & \cdots & S_{1h} \\ \cdots & \cdots & \cdots \\ S_{h1} & \cdots & S_{hh} \end{matrix} \right\|. \tag{18}$$

We have $C_{ii} R_i = S_{ii}$. Here S_{ii} is non-singular and hence C_{ii} has to be non-singular too. The relation (15) defining the anti-automorphic involution now reads as follows:

$$A^* = C^{-1} \bar{A}' C, \qquad CA^* = \bar{A}'C,$$

since in this shape it is invariant with respect to arbitrary, even complex co-ordinate transformations.

We want to prove (*Rosati's theorem*):

If A is even or odd, the matrix C splits up like R corresponding to the numerically different roots α_i of A: $C_{ik} = 0$ for $i \neq k$ in (18). *The roots α are real for an even A; they are pure imaginary for an odd A.*

Indeed by using the even A in its diagonal normal form, the equation

$$A = C^{-1} \bar{A}'C \qquad \text{or} \qquad CA = \bar{A}'C$$

takes on the form

$$C_{ik}(\bar{\alpha}_i - \alpha_k) = 0.$$

Putting $i = k$ yields $\bar{\alpha}_i = \alpha_i$ as C_{ii} is non-singular and hence $\neq 0$. For $i \neq k$ we then get $C_{ik} = 0$ on account of $\alpha_i = \bar{\alpha}_i \neq \alpha_k$. The proof runs along similar lines for an odd A.

The centrum $\mathfrak{z}$ of the algebra $\mathfrak{a}$ is isomorphic to a number field k over ρ of degree h. We are going to replace ρ by k and hence to consider $\mathfrak{a}$ as an algebra over k (the fundamental operations in $\mathfrak{a}$ are then: addition and multiplication of the quantities in $\mathfrak{a}$, multiplication of a quantity in $\mathfrak{a}$ by a number λ in k). Let us apply Rosati's theorem to the matrices A of the centrum only. It proves to be natural, following Lefschetz and Albert, to distinguish two cases.

1) All quantites a of $\mathfrak{z}$ are even: $a^* = a$. By using the determining quantity a_0 of the field $\mathfrak{z}$ and its corresponding matrix A_0, one realizes that k is a totally real field; that is to say, k and all its conjugate fields with respect to ρ are real. In the co-ordinate system in which A_0 is a diagonal matrix, not only R, C, and S, but all matrices A of the algebra $\mathfrak{A}$, break up into parts corresponding to the h numerically different roots α_i of A_0; for every A commutes with A_0. Incidentally, $\rho(\alpha_i) = k_i$ are the h conjugate fields of k. Our problem in ρ reduces to the analogous problem in the "central field" k, the dimensionality n being lowered to $\nu = n/h$.

2) The set $\mathfrak{z}_0$ of the even quantities of $\mathfrak{z}$ does not exhaust the whole $\mathfrak{z}$. Under such circumstances $\mathfrak{z}$ arises from the field $\mathfrak{z}_0$ by adjoining an odd quantity b_0, the square of which b_0^2 lies in $\mathfrak{z}_0$. As $\mathfrak{z}_0$ is isomorphic to a totally real number-field k_0, the centrum $\mathfrak{z}$ is isomorphic to a quadratic extension k of k_0 arising from k_0 by adjoining the square root of a totally negative number γ_0 in k_0. Let us first apply Rosati's theorem to a determining quantity a_0 of $\mathfrak{z}_0$ resulting in the decomposition effected by the transition from ρ to k_0 and then apply it to b_0 for the transition from k_0 to k. The result is the same as in case 1), with the difference, however, that the involution $a \to a^*$ is not reflected as the identity $\lambda \to \lambda$ in the central field k, but as the change $\lambda \to \bar{\lambda}$ to the conjugate complex. The field K into which P extends, by adjoining the numbers of k, coincides with the field of all real numbers in case 1), and with the field of all complex numbers in case 2). The partial matrices R_i of R, (17), lying in K are irreducible in the conjugate fields k_i, for any reduction of them would result in a corresponding reduction of R in ρ. Our problem now has been reduced to the following:

1) Let k be a totally-real number-field, or a number field originating from such a field by adjoining the square root of a totally negative number γ_0 in it. Construct the most general involutorial division algebra over k, with k as its centrum; the involutorial correspondence is supposed to have the properties

$$(a + b)^* = a^* + b^*, \qquad (ab)^* = b^*a^*, \qquad (\lambda a)^* = \bar{\lambda}a^*,$$

where λ is any number of the central field k.

2) Let $\mathfrak{A}$: $a \to A$ be one of the representations of $\mathfrak{a}$ in k (the irreducible one, or one of its multiples), and let C_0 be a symmetric or skew-symmetric non-

$\alpha_1, \cdots, \alpha_h$ all different from each other. By repeating the construction a second time for a vector e' not contained in the partial space thus found, a third time, $\cdots$, finally a ν^{th} time, one breaks up the whole vector space in a number of partial spaces of the described kind. Hence the characteristic polynomial in the total vector space equals $\{\varphi(\lambda)\}^\nu$, and in an appropriate co-ordinate system A becomes a diagonal matrix along the main diagonal of which appears t times the number α_1, then t times the number α_2, and so on. As one readily sees from the equation $RA = AR$, the matrix R splits up in this (non-rational!) co-ordinate system into h matrices of order ν:

$$R = \left\| \begin{matrix} R_1 & & \\ & \ddots & \\ & & R_h \end{matrix} \right\|. \tag{17}$$

Let P be again, and from now on, the field of all real numbers. By performing the transformation into the co-ordinate system just introduced, possibly not real, the matrices C and S shall be treated as the coefficient matrices of bilinear forms of the variables $\bar{x}_\alpha$ and y_α:

$$\textstyle\sum_{\alpha\beta} c_{\alpha\beta}\, \bar{x}_\alpha\, y_\beta\,, \qquad\qquad \sum_{\alpha\beta} s_{\alpha\beta}\, \bar{x}_\alpha\, y_\beta\,,$$

$$C \to \bar{U}'CU\,, \qquad\qquad S \to \bar{U}'SU\,.$$

$\sum s_{\alpha\beta}\,\bar{x}_\alpha\, x_\beta$ then remains a definite Hermitian form; the conditions of symmetry read $\bar{C}' = -C$, $\bar{S}' = S$. In agreement with (17) we write, with reference to the new co-ordinate system,

$$C = \left\| \begin{matrix} C_{11} \cdots C_{1h} \\ \cdots\cdots\cdots \\ C_{h1} \cdots C_{hh} \end{matrix} \right\|, \qquad S = \left\| \begin{matrix} S_{11} \cdots S_{1h} \\ \cdots\cdots\cdots \\ S_{h1} \cdots S_{hh} \end{matrix} \right\|. \tag{18}$$

We have $C_{ii} R_i = S_{ii}$. Here S_{ii} is non-singular and hence C_{ii} has to be non-singular too. The relation (15) defining the anti-automorphic involution now reads as follows:

$$A^* = C^{-1}\bar{A}'\,C\,, \qquad\qquad CA^* = \bar{A}'C\,,$$

since in this shape it is invariant with respect to arbitrary, even complex co-ordinate transformations.

We want to prove (*Rosati's theorem*):

If A is even or odd, the matrix C splits up like R corresponding to the numerically different roots α_i of A: $C_{ik} = 0$ for $i \neq k$ in (18). *The roots α are real for an even A; they are pure imaginary for an odd A.*

Indeed by using the even A in its diagonal normal form, the equation

$$A = C^{-1}\bar{A}'C \qquad \text{or} \qquad CA = \bar{A}'C$$

takes on the form

$$C_{ik}(\bar{\alpha}_i - \alpha_k) = 0\,.$$

Putting $i = k$ yields $\bar{\alpha}_i = \alpha_i$ as C_{ii} is non-singular and hence $\neq 0$. For $i \neq k$ we then get $C_{ik} = 0$ on account of $\alpha_i = \bar{\alpha}_i \neq \alpha_k$. The proof runs along similar lines for an odd A.

The centrum $\mathfrak{z}$ of the algebra $\mathfrak{a}$ is isomorphic to a number field k over ρ of degree h. We are going to replace ρ by k and hence to consider $\mathfrak{a}$ as an algebra over k (the fundamental operations in $\mathfrak{a}$ are then: addition and multiplication of the quantities in $\mathfrak{a}$, multiplication of a quantity in $\mathfrak{a}$ by a number λ in k). Let us apply Rosati's theorem to the matrices A of the centrum only. It proves to be natural, following Lefschetz and Albert, to distinguish two cases.

1) All quantites a of $\mathfrak{z}$ are even: $a^* = a$. By using the determining quantity a_0 of the field $\mathfrak{z}$ and its corresponding matrix A_0, one realizes that k is a totally real field; that is to say, k and all its conjugate fields with respect to ρ are real. In the co-ordinate system in which A_0 is a diagonal matrix, not only R, C, and S, but all matrices A of the algebra $\mathfrak{A}$, break up into parts corresponding to the h numerically different roots α_i of A_0; for every A commutes with A_0. Incidentally, $\rho(\alpha_i) = k_i$ are the h conjugate fields of k. Our problem in ρ reduces to the analogous problem in the "central field" k, the dimensionality n being lowered to $\nu = n/h$.

2) The set $\mathfrak{z}_0$ of the even quantities of $\mathfrak{z}$ does not exhaust the whole $\mathfrak{z}$. Under such circumstances $\mathfrak{z}$ arises from the field $\mathfrak{z}_0$ by adjoining an odd quantity b_0, the square of which b_0^2 lies in $\mathfrak{z}_0$. As $\mathfrak{z}_0$ is isomorphic to a totally real number-field k_0, the centrum $\mathfrak{z}$ is isomorphic to a quadratic extension k of k_0 arising from k_0 by adjoining the square root of a totally negative number γ_0 in k_0. Let us first apply Rosati's theorem to a determining quantity a_0 of $\mathfrak{z}_0$ resulting in the decomposition effected by the transition from ρ to k_0 and then apply it to b_0 for the transition from k_0 to k. The result is the same as in case 1), with the difference, however, that the involution $a \to a^*$ is not reflected as the identity $\lambda \to \lambda$ in the central field k, but as the change $\lambda \to \bar{\lambda}$ to the conjugate complex. The field K into which P extends, by adjoining the numbers of k, coincides with the field of all real numbers in case 1), and with the field of all complex numbers in case 2). The partial matrices R_i of R, (17), lying in K are irreducible in the conjugate fields k_i, for any reduction of them would result in a corresponding reduction of R in ρ. Our problem now has been reduced to the following:

1) Let k be a totally-real number-field, or a number field originating from such a field by adjoining the square root of a totally negative number γ_0 in it. Construct the most general involutorial division algebra over k, with k as its centrum; the involutorial correspondence is supposed to have the properties

$$(a + b)^* = a^* + b^*, \qquad (ab)^* = b^*a^*, \qquad (\lambda a)^* = \bar{\lambda}a^*,$$

where λ is any number of the central field k.

2) Let $\mathfrak{A}$: $a \to A$ be one of the representations of $\mathfrak{a}$ in k (the irreducible one, or one of its multiples), and let C_0 be a symmetric or skew-symmetric non-

singular matrix in k, by means of which the given involution $A \to A^*$ is expressed by the relation

$$A^* = C_0^{-1} \bar{A}' C_0.$$

Construct first the algebra $\mathfrak{L}$ ($\mathfrak{L}_k$ or $\mathfrak{L}_K$) of all matrices in k or K commutable with $\mathfrak{A}$ and then in the most general way a Riemann matrix R (lying in $\mathfrak{L}_K$ and) possessing $\mathfrak{A}$ as its commutator algebra.

§5. Splitting field

The order of the division algebra $\mathfrak{a}$ is a square m^2. With respect to an appropriate splitting field $k(\vartheta)$ over k of degree m, and a corresponding co-ordinate system, the general matrix A of the irreducible representation $a \to A$ of $\mathfrak{a}$ splits up into conjugate m-rowed matrices A_α lying in the m conjugate fields $k(\vartheta_\alpha)$. The individual set $\mathfrak{A}_\alpha = \{A_\alpha\}$ is absolutely irreducible and not only irreducible in $k(\vartheta_\alpha)$. The different A_α are equivalent to each other, since k is the centrum, and accordingly there exist definite non-singular "conjugate" matrices $P_{\alpha\beta}$ in the fields $k(\vartheta_\alpha, \vartheta_\beta)$ satisfying the relations

$$P_{\alpha\beta} A_\beta = A_\alpha P_{\alpha\beta} \qquad (P_{\alpha\alpha} = \text{unit matrix}) \tag{19}$$

for every a. The $P_{\alpha\beta}$ in their turn satisfy equations of the form

$$P_{\alpha\beta} P_{\beta\gamma} = c_{\alpha\beta\gamma} \cdot P_{\alpha\gamma}. \tag{20}$$

The numbers $c_{\alpha\beta\gamma}$ form the *factor set*. With respect to the same co-ordinate system (which is irrational in k), the most general matrix L commutable with all A is of the form

$$\| z_{\alpha\beta} P_{\alpha\beta} \| \qquad (z_{\alpha\beta} \text{ conjugate numbers}).$$

In case 1) $a^* \to A'_\alpha(a)$ as well as $a^* \to A_\alpha(a^*) = A^*_\alpha$, is a representation of $\mathfrak{a}$ in $k(\vartheta_\alpha)$. Hence an equation like

$$A^*_\alpha = C_\alpha^{-1} A'_\alpha C_\alpha \qquad \text{or} \qquad C_\alpha A^*_\alpha = A'_\alpha C_\alpha \tag{21}$$

necessarily holds with fixed non-singular conjugate matrices C_α in $k(\vartheta_\alpha)$. C_α is uniquely determined by this equation but for a numerical factor, as the set of the A_α is absolutely irreducible. Consequently

$$C'_\alpha = \mu_\alpha C_\alpha.$$

This equation leads at once to the condition $\mu_\alpha^2 = 1$ and hence the numerical factor μ_α equals $+1$ or -1. Thus the distinction, C_α symmetric or skew-symmetric, is urged upon us. The matrix C splitting up into the C_α is rational in the original co-ordinate system, and it brings about the transition $A \to A^*$ by means of (15).

The matrices $\check{P}_{\alpha\beta} = P'^{-1}_{\alpha\beta}$, contragredient to the $P_{\alpha\beta}$, satisfy the same relations (19) for A'_α which the $P_{\alpha\beta}$ satisfy for A_α. Hence we must have $\check{P}_{\alpha\beta} = y_{\alpha\beta} P_{\alpha\beta}$

on account of (21). Since the factor set of the $\check{P}_{\alpha\beta}$ equals $1/c_{\alpha\beta\gamma}$ by (20), there follows

$$c^2_{\alpha\beta\gamma} = \frac{y_{\alpha\gamma}}{y_{\alpha\beta}\, y_{\beta\gamma}} \qquad \text{or} \qquad c^2_{\alpha\beta\gamma} \sim 1:$$

the "exponent" of the factor set, and consequently—due to a rather profound proposition concerning division algebras over an algebraic number-field,[6] — the Schur index m must equal 1 or 2.

$m = 1$ is the trivial case where the division algebra $\mathfrak{a}$ over k coincides with k, and where the t-dimensional reducible representation $\mathfrak{A}$ consists of the multiples of the unit matrix lying in k.

In the case $m = 2$, $\mathfrak{a}$ is a quaternion algebra over k; its quantities have the form

$$a = c_0 + c_1\epsilon_1 + c_2\epsilon_2 + c_3\epsilon_3; \qquad (c_i \text{ numbers in } k)$$

$\epsilon_1^2 = e$ and $\epsilon_2^2 = b$ lie in k,[7] $\epsilon_3 = \epsilon_1\epsilon_2 = -\epsilon_2\epsilon_1$. Let us write the quantities a of $\mathfrak{a}$ as

$$a = \rho + \bar{\sigma}\epsilon_2 = (\rho, \sigma)$$

where ρ and σ are numbers in $k(\sqrt{e})$.

$$\rho = c_0 + c_1\epsilon_1 \sim \rho = c_0 + c_1\sqrt{e}, \qquad \bar{\rho} = c_0 - c_1\sqrt{e}; \qquad \sigma = c_2 + c_3\sqrt{e}.$$

The multiplication $(\xi', \eta') = (\rho, \sigma)(\xi, \eta)$ is then expressed as the linear substitution

$$\xi' = \rho\xi + b\bar{\sigma}\eta,$$
$$\eta' = \sigma\xi + \bar{\rho}\eta.$$

Thus the quantities of k can be looked upon as the matrices of the form

$$A_1 = \left\| \begin{matrix} \rho, & b\bar{\sigma} \\ \sigma, & \bar{\rho} \end{matrix} \right\| \qquad \text{in} \qquad k(\sqrt{e}).$$

The irreducible representation of $\mathfrak{a}$ in k if properly normalized, splits up by the substitution

$$(22) \qquad \left\| \begin{matrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \sqrt{e} & 0 & -\sqrt{e} & 0 \\ 0 & \sqrt{e} & 0 & -\sqrt{e} \end{matrix} \right\|$$

[6] R. Brauer, H. Hasse, E. Noether, Journ. f. reine u. angew. Mathem. **167**, 1931, p. 401. Cf. A. A. Albert, Annals of Math. **33**, 1932, pp. 311–318.

[7] e and b are such that the equation $ex^2 + by^2 = 1$ cannot be solved by numbers x, y in k.

into the two conjugate parts A_α ($\alpha = 1, 2$) in the splitting field $k(\sqrt{e})$:

$$A = \left|\begin{array}{cc|cc} \rho, & b\bar{\sigma} & & \\ \sigma, & \bar{\rho} & & \\ \hline & & \bar{\rho}, & b\sigma \\ & & \bar{\sigma}, & \rho \end{array}\right|.$$

From

$$P_{12} = \begin{Vmatrix} 0 & b \\ 1 & 0 \end{Vmatrix}$$

one finds as the most general matrix commuting with all A:

$$L = \left|\begin{array}{cc|cc} \alpha & 0 & 0 & b\beta \\ 0 & \alpha & \beta & 0 \\ \hline 0 & b\bar{\beta} & \bar{\alpha} & 0 \\ \bar{\beta} & 0 & 0 & \bar{\alpha} \end{array}\right|$$

$\alpha, \bar{\alpha}; \beta, \bar{\beta}$ must be two pairs of conjugate numbers in $k(\sqrt{e})$ (or conjugate complex numbers), if this matrix is to lie in k (or the field K of real numbers respectively) after undoing the co-ordinate transformation (22).

Two typical anti-automorphic involutions $a \to a^*$ are the following ones:

$$\epsilon_1 \to -\epsilon_1, \quad \epsilon_2 \to \epsilon_2, \quad \epsilon_3 \to \epsilon_3 \quad \text{or} \quad (\rho, \sigma) \to (\bar{\rho}, \sigma);$$

$$\epsilon_1 \to -\epsilon_1, \quad \epsilon_2 \to -\epsilon_2, \quad \epsilon_3 \to -\epsilon_3 \quad \text{or} \quad (\rho, \sigma) \to (\bar{\rho}, -\sigma).$$

We are going to treat both involutions simultaneously, the upper sign always referring to the first, the lower to the second. A C_α satisfying the equation (21) is given by

$$C_1 = C_2 = \begin{Vmatrix} 0 & 1 \\ \pm 1 & 0 \end{Vmatrix}:$$

$$\begin{Vmatrix} 0, & 1 \\ \pm 1, & 0 \end{Vmatrix} \cdot \begin{Vmatrix} \bar{\rho}, & \pm b\bar{\sigma} \\ \pm\sigma, & \rho \end{Vmatrix} = \begin{Vmatrix} \rho, & \sigma \\ b\bar{\sigma}, & \bar{\rho} \end{Vmatrix} \cdot \begin{Vmatrix} 0, & 1 \\ \pm 1, & 0 \end{Vmatrix}.$$

C_1 is symmetric for the first, and skew-symmetric for the second involution! The matrix C^0 splitting up into C_1 and $C_2 = C_1$ fulfills the equation (15). The most general such matrix has the form C^0L, that is

$$C = \left|\begin{array}{cc|cc} 0 & \alpha & \beta & 0 \\ \pm\alpha & 0 & 0 & \pm b\beta \\ \hline \bar{\beta} & 0 & 0 & \bar{\alpha} \\ 0 & \pm b\bar{\beta} & \pm\bar{\alpha} & 0 \end{array}\right|.$$

C is symmetric for the first involution if α is arbitrary, and $\bar{\beta} = \beta$ (three free parameters in k or K respectively!); it is skew-symmetric if $\alpha = 0, \bar{\beta} = -\beta$ (one parameter). As to the second involution, C is skew-symmetric if α is arbitrary and $\bar{\beta} = -\beta$; it is symmetric if $\alpha = 0, \bar{\beta} = \beta$. This important example shows how the possibilities provided by the general theory may occur together.

The situation is much more complicated in the second case when k is obtained by quadratic imaginary extension of a totally real field. The method of factor sets, however, seems to be appropriate here also for studying the conditions prevailing in a given splitting field of minimum degree m (which need not be either cyclic or even Galois). We note only that the distinction between a symmetric or skew-symmetric C generating the involution $A \to A^* = C^{-1}\bar{A}'C$ now becomes irrelevant. For any symmetric C: $\bar{C}' = C$ gives rise to a skew-symmetric one through multiplication by the purely imaginary number $\sqrt{\gamma_0}$ which extended k_0 to yield k, and vice versa. For all further developments we refer the reader to Albert's paper in the Annals, 1931.

§6. Appendix: Relationship of the new formulation to the usual one

Again we consider the basis α of closed curves on the Riemann surface, and its characteristic form C with coefficients $c_{\alpha\beta}$ $(\alpha, \beta = 1, 2, \cdots, n)$. The matrix C is rational, non-singular and skew-symmetric. Let us choose this time the real basis $dw_\alpha = du_\alpha + idv_\alpha$ of the differentials of the first kind in an arbitrary manner independent of the basis α. The matrices $\| u_{\alpha\beta} \|$, $\| v_{\alpha\beta} \|$ of the real and imaginary parts of the periods

$$u_{\alpha\beta} = \int_\alpha du_\beta, \qquad v_{\alpha\beta} = \int_\alpha dv_\beta$$

may be designated by F and G. From the Dirichlet integral we readily obtain the fact that the bilinear form with the coefficient matrix

$$F'C^{-1}G = S \tag{23}$$

is symmetric and the corresponding quadratic form positive definite. Under the influence of an arbitrary *rational* transformation U of the basis of curves, and an arbitrary *real* transformation V of the basis dw_α of the differentials of the first kind,

$$C; F, G \quad \text{change into} \quad U'CU; U'FV, U'GV, \tag{24}$$

and thus the matrix defined by (23)

$$S \quad \text{into} \quad V'SV.$$

The real matrix $R = \| r_{\alpha\beta} \|$ effecting the transition from dw_α to $-idw_\alpha$:

$$(25) \qquad -idw_\alpha = \sum_\beta r_{\beta\alpha}\, dw_\beta$$

is obtained from the equation

$$(26) \qquad G = FR,$$

and the law by which R changes under the influence of the transformation (24) is given by

$$(27) \qquad R \to V^{-1}RV.$$

We are therefore led to consider the following general situation: C a non-singular skew-symmetric rational matrix, F and G real matrices such that the matrix S defined by (23) is symmetric and "definite." The point is to investigate relations that are invariant with respect to the transformations (24) where U is rational, V real, and both non-singular. In particular, we are concerned with the question of "complex multiplication": how do pairs of non-singular matrices A and B look, A being rational, B real, such that the two equations

$$A'F = FB, \qquad A'G = GB$$

obtain simultaneously? After introducing R by (26), the second equation yields, with respect to the first,

$$(28) \qquad BR = RB.$$

Our way of treatment amounts to the following: by an appropriate transformation V [formula (24), $U = 1$] one takes care that

$$(29) \qquad F \quad \text{becomes} \quad = C' = -C.$$

This equation is preserved under the influence of transformations (24) only if $V = U$. Now S becomes equal to G, and in the problem of complex multiplication (28), $B = C^{-1}A'C$ as well as A is a rational matrix.

The *usual* treatment introduces the assumption $R^2 = -1$. Such an R may be brought into the form

$$(30) \qquad R = \left\| \begin{array}{c|c} 0 & 1 \\ \hline -1 & 0 \end{array} \right\|$$

by an appropriate transformation (27) (the partial squares are p-rowed, $n = 2p$). The substitution (25) is of this form, if the real basis dw_α arises from a complex basis $dw_1, \cdots, dw_p$ in the following manner:

$$dw_1, \cdots, dw_p; \qquad -idw_1, \cdots, -idw_p.$$

The normalization (30) bound to the hypothesis $R^2 = -1$ thus leads to the usual and the normalization (29) requiring no restriction leads to *our* formulation. The greater freedom afforded by the latter for the choice of R should facilitate

considerably the existence proofs,—in particular the proof of the proposition that the problem does not impose any more restrictions on the structure of the commutator algebra than its involutorial character.

The only new idea in this paper is the elimination of the assumption $R^2 = -1$, but I could not avoid retelling the whole story in order to show that this hypothesis is superfluous.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

SYSTEMS OF TOTAL DIFFERENTIAL EQUATIONS DEFINED OVER SIMPLY CONNECTED DOMAINS

By Tracy Yerkes Thomas

(Received May 10, 1934)

1. Consider a system of total differential equations

$$d\varphi^i = \sum_{\alpha=1}^{n} \Psi^i_\alpha (x, \varphi)\, dx^\alpha, \qquad (i = 1, \cdots, M), \tag{1.1}$$

where the functions Ψ are defined over an open simply connected domain D of n dimensions, which can be covered by a system of coordinates x^α, and for arbitrary values of the φ's, i.e. for $-\infty < \varphi^i < +\infty$. All quantities involved are real. As a matter of convenience we shall refer to the domain D of the variables x^α and the range $-\infty < \varphi^i < +\infty$ of the variables φ^i, when considered simultaneously, as the domain Δ.

It is assumed that the functions Ψ are continuous and possess continuous first partial derivatives with respect to the x^α and the φ^i. As an additional assumption, the reason for which will be seen immediately, we impose the condition that the derivatives $\partial\Psi^i_\alpha/\partial\varphi^k$ are bounded in Δ, i.e.

$$\left| \frac{\partial \Psi^i_\alpha}{\partial \varphi^k} \right| < N \tag{1.2}$$

in Δ, where N is a sufficiently large positive constant.

The system of equations

$$\frac{\partial \varphi^i}{\partial x^\alpha} = \Psi^i_\alpha (x, \varphi) \tag{1.3}$$

is completely equivalent to (1.1). Owing to the above assumptions we can form the conditions of integrability of (1.3), namely

$$\frac{\partial \Psi^i_\alpha}{\partial x^\beta} - \frac{\partial \Psi^i_\beta}{\partial x^\alpha} + \sum_{k=1}^{M} \frac{\partial \Psi^i_\alpha}{\partial \varphi^k} \Psi^k_\beta - \sum_{k=1}^{M} \frac{\partial \Psi^i_\beta}{\partial \varphi^k} \Psi^k_\alpha = 0, \tag{1.4}$$

the left members of which are defined in Δ; we assume that (1.4) is satisfied identically in this domain.

In this note we give a simple straightforward proof of the existence theorem for the above system. The method used is in the main similar to that employed by E. Cartan in his book *Géométrie des espaces de Riemann*, 1928, pp. 54–57; the system treated by Cartan is, however, less general than the above system (1.1) and is considered from a point of view peculiar to his special problem.

2. Let P and Q be any two distinct points in D with coordinates x_0^α and x_1^α respectively. Join P and Q by a curve C in D, the curve C being defined by $x^\alpha = f^\alpha(t)$, where the functions f are single valued and have continuous first derivatives; this is possible since D is connected. Suppose $f^\alpha(0) = x_0^\alpha$ and $f^\alpha(1) = x_1^\alpha$ so that t has the range $0 \leqq t \leqq 1$. From (1.3) we obtain

$$\text{(2.1)} \qquad \frac{d\varphi^i}{dt} = \sum_{\alpha=1}^{n} \Psi_\alpha^i\,(f(t), \varphi)\,\frac{df^\alpha}{dt} \equiv \Psi^i\,(t, \varphi),$$

which we regard as equations for the determination of the functions φ^i along C. In consequence of the condition (1.2) and the continuity of the functions Ψ_α^i and f^α we have, for the range $0 \leqq t \leqq 1$ and $-\infty < \varphi^i < +\infty$ of the variables t and φ^i respectively, that (a) the functions Ψ^i are continuous and (b) the derivatives $\partial\Psi^i/\partial\varphi^k$ are bounded. By the theorem of Picard[1] it therefore follows that the equations (2.1) have a solution $\varphi^i(t)$ defined uniquely on C by the arbitrary values $\varphi_0^i = \varphi^i(0)$ of these functions at the point P. The values of the functions $\varphi^i(t)$ at the point Q are then $\varphi^i(1)$. *This proves that if the system* (1.1) *has a solution $\varphi^i(x)$ defined over D which takes on the arbitrarily given values φ_0^i at the point P it can have at most one such solution.* We proceed to prove the existence of such a solution.

3. Join the above points P and Q by another curve C_1 defined parametrically by the equations $x^\alpha = f_1^\alpha(t)$ where $0 \leqq t \leqq 1$, the functions f_1^α being single valued and continuous with continuous first derivatives. We shall show that from the arbitrarily assigned values φ_0^i of the functions φ^i at the point P we shall obtain, by integration of a system of the type (2.1), the same values of the φ^i at Q regardless of whether the integration is carried out with reference to the curve C or the curve C_1.[2]

Since D is simply connected the closed curve formed by C and C_1 can be shrunk to a point. Hence we can pass from C to C_1 by a continuous one parameter family F of such curves, which we represent by

$$x^\alpha = G^\alpha(t, p)\,, \qquad (0 \leqq t \leqq 1,\ 0 \leqq p \leqq 1)\,,$$

such that $x^\alpha = x_0^\alpha$ for $t = 0$ and $x^\alpha = x_1^\alpha$ for $t = 1$ independently of the parameter p. We select functions G^α which possess continuous derivatives

$$\frac{\partial G^\alpha}{\partial t}\,, \qquad \frac{\partial G^\alpha}{\partial p}\,, \qquad \frac{\partial^2 G^\alpha}{\partial t \partial p}\,,$$

---

[1] E. Picard, Traité d'Analyse, Vol. II, 1923, p. 373.

[2] It is sufficient to assume that the curves C and C_1 have no points in common except their end points P and Q and that these curves define different directions at each of their end points. In fact if these conditions are not satisfied by C and C_1 we can take a curve C', analogous to C and C_1, such that the above conditions are satisfied by the pair C and C' and also by the pair C_1 and C'. Then the result that the values of the φ^i at the point Q are the same when determined by integration of (2.1) along C as when determined by integration along C_1 will follow from the corresponding result with reference first to the curves C and C' and second to the curves C_1 and C'.

the second derivatives of the G^α being independent of the order of the differentiation.

Now consider the system

$$(3.1) \qquad \frac{d\varphi^i}{dt} = \sum_{\alpha=1}^{n} \Psi^i_\alpha \, (G(t, p), \varphi) \, \frac{\partial G^\alpha}{\partial t}$$

along any curve of the family F. Taking $\varphi^i = \varphi^i_0$ for $t = 0$, independently of the parameter p, the equations (3.1) determine a unique set of functions $\varphi^i\,(p, t)$ defined for $0 \leqq t \leqq 1$ and $0 \leqq p \leqq 1$. These functions $\varphi^i(p, t)$ are continuous in the variables p and t and possess continuous derivatives $\partial\varphi^i/\partial t$ and $\partial\varphi^i/\partial p$ for all values of the variables for which they are defined. In fact it can be shown that the derivatives $\partial\varphi^i/\partial p$ satisfy the equations[3]

$$\frac{\partial}{\partial t}\left(\frac{\partial\varphi^i}{\partial p}\right) = \sum_{k=1}^{M} \sum_{\alpha=1}^{n} \frac{\partial\Psi^i_\alpha}{\partial\varphi^k} \frac{\partial\varphi^k}{\partial p} \frac{\partial G^\alpha}{\partial t} + \frac{\partial}{\partial p} \sum_{\alpha=1}^{n} \Psi^i_\alpha \frac{\partial G^\alpha}{\partial t}.$$

Hence the derivatives $\partial^2\varphi^i/\partial p\partial t$ which appear in the left members of these equations exist and are continuous; likewise the existence and continuity of the derivatives $\partial^2\varphi^i/\partial t\partial p$ result by differentiation of the equations (3.1). Hence it follows that[4]

$$(3.2) \qquad \frac{\partial^2\varphi^i}{\partial p\partial t} = \frac{\partial^2\varphi^i}{\partial t\partial p},$$

for $0 \leqq t \leqq 1$ and $0 \leqq p \leqq 1$, i.e. these second derivatives are independent of the order of differentiation.

Now form the equations

$$(3.3) \qquad \frac{\partial\varphi^i}{\partial t} = \sum_{\alpha=1}^{n} \Psi^i_\alpha \frac{\partial G^\alpha}{\partial t},$$

$$(3.4) \qquad \frac{\partial\varphi^i}{\partial t} = \sum_{\alpha=1}^{n} \Psi^i_\alpha \frac{\partial G^\alpha}{\partial p} + \sigma^i,$$

the functions σ^i being defined by these latter equations. Then by differentiation of (3.3) and (3.4) we obtain

---

[3] See, Frank-v. Mises, *Differentialgleichungen der Physik*, 2nd Ed., Braumschweig 1930, p. 287, where these results are proved for the case of a single equation. The extension to a system of equations, such as the above equations (3.1), is immediate.

[4] See, for example, E. Goursat, Cours d'Analyse Mathématique, 5th Ed., Paris, 1927, p. 42.

$$\frac{\partial^2\varphi^i}{\partial t\partial p} = \sum_{k=1}^{M}\left(\sum_{\alpha=1}^{n}\frac{\partial\Psi^i_\alpha}{\partial\varphi^k}\frac{\partial x^\alpha}{\partial t}\right)\left(\sum_{\beta=1}^{n}\Psi^k_\beta\frac{\partial x^\beta}{\partial p}+\sigma^k\right)$$

$$+\sum_{\alpha=1}^{n}\Psi^i_\alpha\frac{\partial^2 x^\alpha}{\partial t\partial p}+\sum_{\alpha,\beta=1}^{n}\frac{\partial\Psi^i_\alpha}{\partial x^\beta}\frac{\partial x^\alpha}{\partial t}\frac{\partial x^\beta}{\partial p},$$

$$\frac{\partial^2\varphi^i}{\partial p\partial t} = \sum_{\alpha=1}^{n}\Psi^i_\alpha\frac{\partial^2 x^\alpha}{\partial p\partial t}+\frac{\partial\sigma^i}{\partial t}+\sum_{\alpha,\beta=1}^{n}\frac{\partial\Psi^i_\alpha}{\partial x^\beta}\frac{\partial x^\alpha}{\partial p}\frac{\partial x^\beta}{\partial t}$$

$$+\sum_{k=1}^{M}\sum_{\alpha,\beta=1}^{n}\frac{\partial\Psi^i_\alpha}{\partial\varphi^k}\Psi^k_\beta\frac{\partial x^\alpha}{\partial p}\frac{\partial x^\beta}{\partial t}.$$

Subtracting corresponding members of these equations leads, on account of (3.2), to the set of equations

$$\frac{\partial\sigma^i}{\partial t} = \sum_{k=1}^{M}\sum_{\alpha=1}^{n}\frac{\partial\Psi^i_\alpha}{\partial\varphi^k}\frac{\partial x^\alpha}{\partial t}\sigma^k, \tag{3.5}$$

when use is made of the integrability conditions (1.4).

Since $\varphi^i = \varphi^i_0$ and $x^\alpha = x^\alpha_0$ for $t = 0$, independently of p, we have from (3.4) that $\sigma^i = 0$ for $t = 0$. Hence from (3.5) it follows that $\sigma^i = 0$ along any curve of parameter p so that, in particular, $\sigma^i = 0$ for $t = 1$. Then from (3.4) the derivatives $\partial\varphi^i/\partial p = 0$ for $t = 1$. Hence the value of φ^i at the point Q is independent of the curve of integration of the family F by which the point P is joined to the point Q. *In other words there exists a set of functions $\varphi^i(x)$ defined throughout the domain D, these functions being uniquely determined by the assignment of their values φ^i_0 at an arbitrary point P and the process of integration of the system* (2.1) *along curves of the type C issuing from P.*

4. Now consider the values $\varphi^i(\bar{x})$ of the above functions $\varphi^i(x)$ at an arbitrary point $\bar{P}$ of the domain D. Starting with the point $\bar{P}$ and the values $\varphi^i(\bar{x})$ we can determine by the above process, i.e. by integration of (2.1) along curves C issuing from P, a set of functions $\bar{\varphi}^i(x)$ analogous to the functions $\varphi^i(x)$. It is then evident that the functions $\bar{\varphi}^i(x)$ will take the values of the corresponding functions $\varphi^i(x)$ at the point P used in the determination of these latter functions; also to determine the values of the functions $\bar{\varphi}^i$ at any point Q of D we can integrate (2.1) along a curve C which passes through the point P. *Hence the functions $\bar{\varphi}^i(x)$ at which we arrive, when we start with any other point $\bar{P}$ of the domain D in the process of the determination of these functions, are identical with the functions $\varphi^i(x)$ provided that the values of the functions $\bar{\varphi}^i(x)$ at the point $\bar{P}$ are the same as the values of the corresponding functions $\varphi^i(x)$ at this point.*

It remains to observe that the above functions $\varphi^i(x)$ satisfy the system (1.1). But this is seen immediately. In fact let Q be any point of the domain D and take the curve C as the curve of parameter x^α passing through Q. Then, in

view of the last italicized result, the equations (2.1) are satisfied along this curve C by the functions $\varphi^i(x)$; this gives

$$\frac{\partial \varphi^i(x)}{\partial x^\alpha} = \Psi^i_\alpha\,(x, \varphi(x))$$

at all points of the domain D.

THEOREM. *Let D be an open simply connected domain of n dimensions which can be covered by a system of coordinates x^α and denote by Δ the domain composed of D and the range $-\infty < \varphi^i < +\infty$ simultaneously. Suppose that the functions $\Psi^i_\alpha(x, \varphi)$ defined in Δ are continuous and possess continuous first partial derivatives in this domain; suppose, furthermore, that the equations* (1.4) *are satisfied identically and that the derivatives $\partial\Psi^i_\alpha/\partial\varphi^k$ are bounded in Δ. Then the system* (1.1) *admits a unique solution $\varphi^i(x)$ defined throughout the domain D such that the functions $\varphi^i(x)$ take an arbitrary set of values φ^i_0 at an arbitrary point P of D.*

5. In particular the domain D may be homeomorphic to the interior of an n dimensional Euclidean hypersphere Σ; this is the simplest and undoubtedly most important special case of the above theorem. Going out from this special case we can extend the Theorem to open or closed simply connected spaces S each point of which belongs to a neighborhood capable of being put into one to one reciprocal correspondence with the interior of Σ.[5] Such a space S can therefore be covered by one or more coordinate systems. Assuming the scalar character of the functions φ^i, and the property of differentiability of the coordinate relations throughout portions of S common to two coordinate systems, it is evident that the above discussion will continue to apply since the equations involved are invariant under coordinate transformations. In the statement of the above theorem the domain D can accordingly be replaced by the space S.

PRINCETON UNIVERSITY.

[5] For a precise characterization of the space S see E. Cartan, *La théorie des groupes finis et continus et l'Analysis Situs*, Mém. des Sciences. Math. No. 42, Gauthier-Villars, 1930, p. 3. In the postulates for the space S closed neighborhoods are used by Cartan. The corresponding postulates involving open neighborhoods are given by T. Y. Thomas, *The Differential Invariants of Generalized Spaces*, Cambridge University Press, 1934, p. 1.

ANNALS OF MATHEMATICS
Vol. 35, No. 4, October, 1934

SIMPLE TENSORS AND THE PROBLEM OF THE INVARIANT CHARACTERIZATION OF AN N-TUPLY ORTHOGONAL SYSTEM OF HYPERSURFACES IN A V_n

BY TRACY YERKES THOMAS AND JACK LEVINE

1. Let V_n be a metric space with fundamental quadratic differential form

$$\Phi = g_{\alpha\beta}(x)\, dx^\alpha\, dx^\beta\,, \tag{1}$$

and denote by T a tensor[1] differential invariant of V_n. The tensor T will be called a *simple tensor* if its components possess the following four properties:

(a) They are polynomials in the $g_{\alpha\beta}$, the $g^{\alpha\beta}$ and their derivatives,

(b) they are linear in the derivatives of highest order and these derivatives have coefficients which are functions of the $g_{\alpha\beta}$,

(c) they can be written as expressions valid for an infinite set S of values of the dimensional number n,

(d) the numerical coefficients of the $g_{\alpha\beta}$ and their derivatives enter as rational functions of n.

We observe that the above definition of the simple tensor can be extended to tensor differential invariants of spaces of distant parallelism, affine spaces, etc., as well as the affine representations of conformal and projective spaces. Such tensors are obtained in the theory of differential invariants as the simplest type of conditions of integrability. The simple tensors include all the important differential invariants of the above spaces. In fact, in all such cases, simple tensors suffice for the solution of the equivalence problem, i.e. constitute a complete set of differential invariants.[2]

2. The problem of determining the existence of an n-tuply orthogonal system of hypersurfaces in a V_n is equivalent to that of finding the conditions under which the form (1) is reducible to the sum of squares, i.e. to the form

$$\sum_{\alpha=1}^{n} h_{\alpha\alpha}(y)\, dy^\alpha\, dy^\alpha \tag{2}$$

by a suitable coordinate transformation. As is well known a *sufficient* condition for the existence of the orthogonal hypersurfaces is furnished by the vanishing

---

[1] The term tensor as here used is to be regarded as including the scalar.

[2] It is to be doubted however if simple tensors exist which constitute a complete set of differential invariants of the projective or conformal spaces when reference is not had to the affine representations of these spaces. The question of the existence or non-existence of a complete set of tensor differential invariants of the n-dimensional projective and conformal spaces gives rise to an interesting problem which has not yet been solved.

of the Riemannian curvature tensor;[3] likewise by the vanishing of the conformal curvature tensor.[4]

It is to be observed that the above tensors whose vanishing give sufficient conditions for the reduction of (1) to (2) are simple tensors. When first treating the problem of finding *necessary and sufficient* conditions of tensor character for the reduction of (1) to the form (2) it is therefore natural to attempt to do this by means of simple tensors. In connection with this attempt we shall need a lemma which will now be proved.

It is possible to eliminate all derivatives of $g^{\alpha\beta}$ so as to have only derivatives of the $g_{\alpha\beta}$ appearing in the components of a tensor differential invariant. The former derivatives may in fact be replaced by derivatives of the $g_{\alpha\beta}$ by means of the relations

$$\tag{3} \begin{gathered} \frac{\partial g^{\mu\nu}}{\partial x^\rho} = -\, g^{\mu\alpha} g^{\nu\beta} \frac{\partial g_{\alpha\beta}}{\partial x^\rho}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial^r g^{\mu\nu}}{\partial x^\rho \cdots \partial x^\sigma} = -\, g^{\mu\alpha} g^{\nu\beta} \frac{\partial^r g_{\alpha\beta}}{\partial x^\rho \cdots \partial x^\sigma} + \star, \end{gathered}$$

where the $\star$ denotes polynomials in derivatives of $g_{\alpha\beta}$ of lower order than r. We shall hereafter suppose that all derivatives of the $g^{\alpha\beta}$ have been eliminated by means of the above relations in the components of the tensors under consideration.

LEMMA. *The set of coefficients of the derivatives of highest order of a tensor differential invariant which is linear in these derivatives, constitutes the components of a tensor differential invariant.*

Letting T denote such a tensor we can suppose without loss of generality that only covariant indices are involved in the symbol of its components. Then the components of T can be written

$$\tag{4} T_{\alpha\cdots\beta} = M^{\mu\nu\sigma\cdots\tau}_{\alpha\cdots\beta} \cdot \frac{\partial^r g_{\mu\nu}}{\partial x^\sigma \cdots \partial x^\tau} + \star,$$

where the above coefficients M are symmetric in the indices $\mu\nu$ and also in the indices $\sigma \cdots \tau$. It is required to prove that the set of quantities M are the

[3] In fact, the vanishing of this tensor is necessary and sufficient for the form (1) to be reducible to a form (2) with constant coefficients.

[4] This tensor is sometimes called the Weyl conformal curvature tensor and is defined for $n \geqq 4$; its vanishing is a necessary and sufficient condition for the reduction of the form (1) to a form

$$H(y^1, \cdots, y^n) \cdot \sum_{\alpha=1}^{n} c_\alpha \, dy^\alpha \, dy^\alpha$$

where $c_\alpha = \pm 1$. The orthogonal hypersurfaces always exist for $n = 2, 3$.

components of a tensor; we observe that the M's are functions of the $g_{\alpha\beta}$ and their derivatives to the order $r - 1$ inclusive. Now consider the transformation equations

$$\bar{T}_{\gamma \cdots \delta} = T_{\alpha \cdots \beta} \frac{\partial x^\alpha}{\partial \bar{x}^\gamma} \cdots \frac{\partial x^\beta}{\partial \bar{x}^\delta}, \tag{5}$$

where the components $T_{\alpha \cdots \beta}$ are given by the right members of (4) and the components $\bar{T}_{\gamma \cdots \delta}$ by corresponding expressions in the barred quantities. Eliminate the $g_{\alpha\beta}$ from the components $T_{\alpha \cdots \beta}$ in (5) by the substitution

$$g_{\alpha\beta} = \bar{g}_{\gamma\delta} \cdot \frac{\partial \bar{x}^\gamma}{\partial x^\alpha} \cdot \frac{\partial \bar{x}^\delta}{\partial x^\beta},$$

also eliminate the derivatives of the $g_{\alpha\beta}$ by

$$\frac{\partial^s g_{\alpha\beta}}{\partial x^\mu \cdots \partial x^\nu} = \frac{\partial^s \bar{g}_{\gamma\delta}}{\partial \bar{x}^\pi \cdots \partial \bar{x}^\rho} \cdot \frac{\partial \bar{x}^\gamma}{\partial x^\alpha} \cdots \frac{\partial \bar{x}^\rho}{\partial x^\nu} + \star .$$

We then obtain

$$\bar{M}^{c\, d\, \mu \cdots \nu}_{\gamma \cdots \delta} \cdot \frac{\partial^r \bar{g}_{cd}}{\partial \bar{x}^\mu \cdots \partial \bar{x}^\nu} + \star = \left(M^{a\, b\, \sigma \cdots \tau}_{\alpha \cdots \beta} \cdot \frac{\partial^r g_{ab}}{\partial x^\sigma \cdots \partial x^\tau} + \star \right) \cdot \frac{\partial x^\alpha}{\partial \bar{x}^\gamma} \cdots \frac{\partial x^\beta}{\partial \bar{x}^\delta}$$

$$= M^{a\, b\, \sigma \cdots \tau}_{\alpha \cdots \beta} \cdot \frac{\partial^r \bar{g}_{cd}}{\partial \bar{x}^\mu \cdots \partial \bar{x}^\nu} \cdot \frac{\partial x^\alpha}{\partial \bar{x}^\gamma} \cdots \frac{\partial x^\beta}{\partial \bar{x}^\delta} \cdot \frac{\partial \bar{x}^c}{\partial x^a} \cdots \frac{\partial \bar{x}^\nu}{\partial x^\tau} + \star$$

Since the third member of these equations, considered as expressions in the barred quantities, is identically equal to the first member, it follows that

$$\bar{M}^{c\, d\, \mu \cdots \nu}_{\gamma \cdots \delta} = M^{a\, b\, \sigma \cdots \tau}_{\alpha \cdots \beta} \cdot \frac{\partial x^\alpha}{\partial \bar{x}^\gamma} \cdots \frac{\partial \bar{x}^\nu}{\partial x^\tau},$$

and hence the M's are the components of a tensor.

It follows as a corollary of the above lemma that the set of coefficients of the highest ordered derivatives of a simple tensor, are the components of a tensor.

3. Suppose now we have a simple tensor T which we can take without loss of generality to be completely covariant, i.e. to have components which can be written in the form (4) and such that its vanishing constitutes a *necessary* condition for the reduction of (1) to the form (2). The coefficients M in (4) are now functions of the $g_{\alpha\beta}$ and constitute the components of a tensor by the above lemma. Since $T = 0$ when the form (1) can be reduced to (2), we have

$$N^{a\, a\, \alpha \cdots \beta}_{\sigma \cdots \tau} \cdot \frac{\partial^r h_{aa}}{\partial y^\alpha \cdots \partial y^\beta} + \star = 0$$

identically in the $h_{\alpha\alpha}$ and their derivatives where the N's are obtained from the M's in (4) by transformation to the y coordinate system. Hence we must have

$$N^{a\, a\, \alpha \cdots \beta}_{\sigma \cdots \tau} (h_{ii}) = 0 , \qquad (a \text{ not summed}), \tag{6}$$

identically in the n quantities h_{ii}.

To simplify the following discussion we reduce the components to a completely contravariant form and thus obtain the set of components

$$N^{a\,b\,\alpha\cdots\beta\gamma\cdots\delta} = N^{a\,b\,\alpha\cdots\beta}_{\sigma\cdots\tau} \cdot h^{\sigma\gamma} \cdots h^{\tau\delta}\,. \tag{7}$$

In terms of these latter components the equations (6) become

$$N^{a\,a\,\alpha\beta\gamma\delta\cdots}(h^{ii}) = 0\,, \qquad (a \text{ not summed}) \tag{8}$$

where it is to be noticed that the left members of these equations involve only the contravariant components h^{ii}.

We shall show that the left members of (7) vanish identically whence *it will follow that there exists no simple tensor whose vanishing is a necessary condition for the reduction of* (1) *to the form* (2).

The most general form of a set of tensor components of the type occurring in the left members of (7), these components depending only on the h^{ij}, is given by a sum of terms[5]

$$\sum C \cdot h^{ab} \cdot h^{\alpha\beta} \cdots h^{\gamma\delta}, \tag{9}$$

where we use all possible permutations of the indices $ab\alpha\beta \cdots \gamma\delta$ which give distinct terms and each term thus obtained is multiplied by an arbitrary constant C. In consequence however of the above properties characterizing the simple tensor, the coefficients C for the tensor under consideration are in fact rational functions $C(n)$ of the dimensional number n of the space. Since the components N in (7) are symmetric in the first two indices ab, we can write

$$\begin{aligned} N^{a\,b\,\alpha\beta\gamma\delta\cdots} = h^{a\,b}\, N_1^{\alpha\beta\gamma\delta\cdots} &+ (h^{a\alpha}\, h^{b\beta} + h^{a\beta}\, h^{b\alpha})\, N_2^{\gamma\delta\cdots} \\ &+ (h^{a\alpha}\, h^{b\gamma} + h^{a\gamma}\, h^{b\alpha})\, N_3^{\beta\delta\cdots} + \cdots. \end{aligned} \tag{10}$$

Now in (10) put $h^{ij} = 0$ $(i \neq j)$ and take $a = b \neq \alpha, \beta, \cdots$. Then by (8) we have

$$N_1^{\alpha\beta\gamma\delta\cdots}(h^{ii}) = 0\,, \qquad (\alpha, \beta, \gamma, \delta, \cdots \neq a). \tag{11}$$

Now suppose that there are $2k$ indices in the set $\alpha\beta\gamma\delta \cdots$ and assume $n > 2k$. Let any term in the sum of terms which comprise the quantities N_1 in (10) be denoted by

$$C(n) \cdot h^{\mu\nu}\, h^{\sigma\tau} \cdots h^{\lambda\rho}, \tag{12}$$

where $\mu\nu\sigma\tau \cdots \lambda\rho$ denotes a permutation of the indices of the set $\alpha\beta\gamma\delta \cdots$. Take $\mu = \nu, \sigma = \tau, \cdots, \lambda = \rho$ such that no two pairs of these indices have the same value and no pair has a value equal to that of the index a. Hence all terms in (11) become zero except the term (12) with the result that the coefficient $C(n)$ in (12) is equal to zero. In other words the quantities N_1 in (10) vanish identically.

---

[5] The corresponding theorem for completely covariant tensors appears as a special case of a theorem given by P. Franklin, Tensors of given Types in Riemann Space, Phil. Mag., vol. 45, (1923), p. 1009. From this result the form (9) is immediately obtained.

We now take $a = b = \alpha = \beta \neq \gamma, \delta, \cdots$, and obtain

$$N_{2}^{\gamma\delta\cdots}(h^{ii}) = 0, \qquad (\gamma, \delta, \cdots, \neq a). \tag{13}$$

By proceeding as for (11) it can be shown that the quantities N_2 in the right members of (10) are equal to zero identically. Continuing, we arrive at the conclusion that all components N in the left members of (10) vanish identically if $n > 2k$, i.e., all coefficients $C(n)$ in (10) equal zero for $n > 2k$. Since the $C(n)$ are rational functions of n by hypotheses, they must therefore vanish for all values of n in the set S under consideration.

PRINCETON UNIVERSITY.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

THE COMPLETE INDEPENDENCE OF CERTAIN PROPERTIES OF MEANS

By E. L. Dodd

(Received September 28, 1933)

§1. Introduction

The mean m of a finite or countable set of numbers $x_1, x_2, \cdots$, will be taken as

$$m = F(x_1, x_2, \cdots), \tag{1}$$

with but a single condition placed upon the function F, namely: For each x,

$$F(x, x, \cdots) = x, \tag{2}$$

uniquely,—although $F(x_1, x_2, \cdots)$ may in general be multiple-valued. This condition (2)—for a finite set of numbers—was used by O. Suto[1] as one of three conditions to characterize the arithmetic mean; and it has been used by several other writers.

It will not be required that the mean m be intermediate between the least and the greatest of the x's—or of their lower and upper bounds. Indeed, O. Chisini[2] and B. de Finetti[3] have given illustrations of "external" means. And the mean in Illustration 3, which I give in a later section of this paper, is of this sort. Intermediacy or internality ceases to be of great importance if emphasis is laid upon the *use* to which the mean is to be put. If the condition of intermediacy is satisfied, then (2) is satisfied; but the converse of this statement is not true.

For the theorems of this paper, the uniqueness of $F = x$ in (2) will be adhered to strictly; but a useful extension of the definition of a mean—which merely requires that at least *one* value of $F(x, x, \cdots, x)$ shall be x—will be exemplified in the first part of Illustration 5.

Definitions of certain properties of means—following De Finetti—will now be given for means of a finite or countable set of elements. The accompanying abbreviations will be used subsequently either for the adjective or for the corresponding noun. The symbol Max will stand for the maximum element or the least upper bound of the elements, and Min for the least element or the greatest lower bound of the elements.

---

[1] Law of the arithmetic mean. Tohoku Math. Journ., vol. 5 (1914), pp. 79–81.

[2] Sul concetto di media, Periodico di Matematico, ser. 4, vol. 9 (1929), pp. 106–116.

[3] Sul concetto di media, Giornale dell' Istituto Italiano degli Attuari, vol. 2 (1931), pp. 369–396.

Definitions

Un. A mean is *unique*, if it is single-valued.

Int. A mean m is *internal* if Min $(x_1, x_2, \cdots) \leqq m \leqq$ Max $(x_1, x_2, \cdots)$.

Mon. A mean $m = f(x_1, x_2, \cdots)$ is monotone if $f(x'_1, x'_2, \cdots) \geqq f(x_1, x_2, \cdots)$ when each $x'_i \geqq x_i$.

Inc. A mean m is *increasing* if $f(x'_1, x'_2, \cdots) > f(x_1, x_2, \cdots)$ when each $x'_i \geqq x_i$ and some $x'_r > x_r$.

Sym. A mean is *symmetrical* if the interchange of any two elements (at pleasure) does not alter the mean.

Hom. A mean m is *homogeneous* if $f(kx_1, kx_2, \cdots) = kf(x_1, x_2, \cdots)$, where k is a constant, either positive, negative, or zero.

Tr. A mean m is *translative* if $f(x_1 + k, x_2 + k, \cdots) = f(x_1, x_2, \cdots) + k$, k = constant.

As. A mean is *associative* if it remains unchanged when in any subset of its elements (taken at pleasure) each element of the subset is replaced by the mean for that subset.

M. Nagumo[4] proved that an associative mean m of n numbers, $x_1, x_2, \cdots, x_n$, has the form

$$m = \psi^{-1}\left[\frac{1}{n}\sum \psi(x_i)\right], \tag{3}$$

where ψ is continuous and increasing—with inverse ψ^{-1}—provided that the mean is symmetrical, continuous, internal, and (2) is satisfied.

The foregoing definitions will be given a strict interpretation. Thus, the weighted mean $(c_1x_1 + c_2x_2 + c_3x_3)/(c_1 + c_2 + c_2)$ is not symmetric when the weights c_i are different. Nor is this mean associative. For if x_2 and x_3 are each replaced by $(k_1x_2 + k_2x_3)/(k_1 + k_2)$, the mean in general will be changed. In case a mean is multiple-valued, it does not have a stated property unless each branch has that property.

The graphical interpretation of the associative mean, when given by (3), is important. From the points x_i on the X axis, ordinates are drawn to the curve $y = \psi(x)$. From these points of intersection, (x_i, y_i), horizontal lines are drawn to the Y axis. For the points on the Y axis thus obtained, the point $y' = \Sigma y_i/n$ representing the *arithmetic mean* of the y_i's is obtained. From this point a horizontal line is drawn to the point (x', y') on the curve; and from (x', y') a vertical line to $(x', 0)$ on the X axis. The mean $m = x'$.

§2. Complete Independence of Certain Properties

Simple independence for several properties was shown in the papers mentioned. Complete independence of various sets of properties will now be shown. Complete independence of r properties is shown when 2^r means are indicated,

---

[4] Über eine Klasse der Mittelwerte, Japanese Journ. of Math., vol. 7 (1930), pp. 71–79.

one for each combination of the properties, in presence or in absence. In Tables I and II, the means are continuous; in Table III, discontinuous. All the means considered are defined for all real values of the n elements, $x_1, x_2, \cdots, x_n$; indeed, on each branch when the mean is multiple valued.

The means used in Tables I and II are simple combinations of the minimum (Min), the maximum (Max), the arithmetic mean, A; a power mean, P; the exponential mean, E; or their weighted forms. Summation from 1 to n will be indicated by Σ.

$$
\begin{array}{llll}
 & A = \sum x_i/n; & A' = \sum c_i x_i/c; & c_i > 0, \quad \sum c_i = c; \\
(4) & P = [\sum x_i^3/n]^{1/3}; & P' = [\sum c_i x_i^3/c]^{1/3}; & c_i\text{'s not all equal}; \\
 & E = \log_b[\sum b^{x_i}/n]; & E' = \log_b[\sum c_i b^{x_i}/c]; & b > 0, \quad b \neq 1.
\end{array}
$$

In the tables, the absence of a property is indicated by a dash.

TABLE 1

UNIQUE, INTERNAL, SYMMETRIC MEANS

$\alpha > 0, \quad \beta > 0, \quad \alpha + \beta = 1$

Increasing			Not Increasing		
$\alpha A + \beta(\text{Min} + \text{Max})/2$	Hom	Tr	$\text{Min} + \text{Max} - \alpha A - \beta(\text{Min} + \text{Max})/2$	Hom	Tr
$\alpha A + \beta P$	Hom	—	$\text{Min} + \text{Max} - \alpha A - \beta P$	Hom	—
$\alpha A + \beta E$	—	Tr	$\text{Min} + \text{Max} - \alpha A - \beta E$	—	Tr
$\alpha P + \beta E$	—	—	$\text{Min} + \text{Max} - \alpha P - \beta E$	—	—

It is obvious that if I and I' are Int. means, so also is $\alpha I + \beta I' \leqq \alpha$ Max $+ \beta$Max $=$ Max. So also is Min $+$ Max $- I$; since Min $- I \leqq 0$, and Max $- I \geqq 0$. Moreover, if I is Inc, Min $+$ Max $- I$ is not Inc.; for it does not increase when an element, not an extreme, is given a small increase. It is supposed known that A is Hom and Tr; P is Hom and not Tr; E is Tr but not Hom. By indirect proof, it can be shown that a combination like $\alpha A + \beta P$ —of a Tr and non-Tr mean—is non-Tr. Likewise, for Hom.

Now let A, P, and E be primed. The foregoing means become asymmetric; but otherwise, they have the same properties.

It follows that Inc, Hom, Tr, and Sym are completely independent.

Now let $m_1 = f_1(x_1, x_2, \cdots, x_n)$ be any mean in Table I, and let $m_2 = f_2$ be the result of replacing α and β by α' and β' with $\alpha' > 0, \beta' > 0, \alpha' + \beta' = 1$, $\alpha' \neq \alpha$. Take $m = m_1$ or m_2 that is, take

$$m = F(x_1, x_2, \cdots, x_n) = (f_1 + f_2)/2 \pm (f_1 - f_2)/2.$$

Then, since f_1 and f_2 satisfy (2), it follows that $F(x, x, \cdots, x) = x$ uniquely, altho F is in general two-valued. Each branch, m_1 and m_2, considered *individually,* has the properties designated in Table I. This proves

THEOREM I. *For means $m = F(x_1, x_2, \cdots, x_n)$, continuous and defined for all real values of $x_1, x_2, \cdots, x_n$, where for each x,*

$$F(x, x, \cdots x) = x \tag{2}$$

uniquely,—although $F(x_1, x_2, \cdots x_n)$ may in general be double-valued, the properties of being increasing, unique, homogeneous, translative, and symmetric, are completely independent.

In place of increasing we may substitute monotone, for the means listed as non-increasing are not even monotone increasing.

The following is a table for non-internal means, i.e., means not internal for all values of the elements.

TABLE II

UNIQUE, NON-INTERNAL, SYMMETRIC MEANS

$(\text{Min} + \text{Max})/2 + \gamma\,[A - (\text{Min} + \text{Max})/2]$	Hom	Tr
$(\text{Min} + \text{Max})/2 + \gamma\,[P - (\text{Min} + \text{Max})/2]$	Hom	—
$(\text{Min} + \text{Max})/2 + \gamma\,[E - (\text{Min} + \text{Max})/2]$	—	Tr
$(\text{Min} + \text{Max})/2 + \gamma\,[P + E - \text{Min} - \text{Max}]/2$	—	—

These are single-valued means satisfying $F(x, x, \cdots, x) = x$.

The [ ] above is not in general equal to zero. Then with γ sufficiently large, the above means are external. E.g., with $\gamma = 2$, take $x_1 = \cdots = x_9 = 1$; $x_{10} = 2$; $b = 2$. The means are Sym and the priming of A, P, and E leads to Asym means with other properties unaltered. These means may be transformed into double-valued means by using γ and γ', with $\gamma' \neq \gamma$. Tables I and II lead to

THEOREM II. *For means $m = F(x_1, x_2, \cdots x_n)$, continuous and defined for all real values of $x_1, x_2, \cdots, x_n$, where for each x,*

$$F(x, x, \cdots x) = x \tag{2}$$

uniquely,—although $F(x_1, x_2, \cdots x_n)$ may in general be double-valued, the properties of being internal, unique, homogeneous, translative, and symmetric, are completely independent.

In Table III of discontinuous means, the first mean, with $\kappa = 1$, is due to R. D. Beetle[5] (p.4 46). A = Arithmetic mean. Sgn $x = -1, 0$, or $+1$, according as $x < 0$, $x = 0$, or $x > 0$. Summation, $\sum$, for $i = 1, 2, \cdots, n$; $j = 1, 2, \cdots, n$.

[5] On the complete independence of Schimmack's postulates for the arithmetic mean, Math. Ann., vol. 76 (1915), pp. 444–446.

TABLE III

UNIQUE, DISCONTINUOUS MEANS

$\kappa > 0 < \lambda < 1/4n$

$A + \kappa \operatorname{sgn} (x_1 - x_n)$	—	Tr	—	—
$A + \kappa \operatorname{sgn} \sum (x_i - x_j)^2$	—	Tr	Sym	—
$A + A\kappa \operatorname{sgn} \sum (x_i - x_j)^2$	Hom	—	Sym	—
$A + A\kappa \operatorname{sgn}(x_1 - x_n)$	—	—	—	—
$A + \lambda(x_1 - x_n)[\operatorname{sgn}(x_1^2 - x_n^2) + \operatorname{sgn}(x_1^2 - x_n^2)^2]$	Hom	—	—	Inc
$A + \lambda(x_1 - x_n)[\operatorname{sgn}(x_1 - x_n) + \operatorname{sgn}(x_1 - x_n)^2][\operatorname{sgn}(x_1 + x_n) + \operatorname{sgn}(x_1 + x_n)^2]$		—	—	— Inc

The fifth and sixth means are obviously increasing, because of the condition $0 < \lambda < 1/4n$, except possibly at points of discontinuity given by $x_1 + x_n = 0$. The factor $(x_1 - x_n)$ prevents a discontinuity at $x_1 - x_n = 0$. The fifth mean equals A, unless $x_1^2 - x_n^2 = (x_1 - x_n)(x_1 + x_n) > 0$. Starting with $x_1 > 0$, $x_n = -x_1$, if either x_1 or x_n is increased algebraically, the mean is given a positive jump. Again, starting with $x_1 < 0, x_n = -x_1$, if either x_1 or x_n is decreased algebraically, the mean is given a negative jump. In the sixth mean, the discontinuity arises only for $x_1 + x_n = 0$, with $x_1 > 0$; and homogeneity is lost.

The means in Table III are unique; but from them we may construct multiple-valued means by using different values of κ or λ.

From Tables I and III, the following theorem may be inferred:

THEOREM III. *For means $m = F(x_1, x_2, \cdots x_n)$, defined for all real values of $x_1, x_2, \cdots, x_n$, where for each x,*

$$F(x, x, \cdots, x) = x \tag{2}$$

uniquely,—although $F(x_1, x_2, \cdots x_n)$ may in general be double-valued, there is complete independence for three sets of properties, each set including continuity and uniqueness, the other properties being: Translativeness and symmetry; Homogeneity and symmetry; and Homogeneity and increasing univariance.

It is known that A, P, and E are As means, and it is easy to form a mean—see equation (7) with the c_r's all equal—which is neither Hom nor Tr. Hence, it will be regarded as known that As, Hom, and Tr are completely independent.

§3. Dependence of Certain Properties

If a mean is As, then R. Schimmack's[6] fourth condition is satisfied. Hence a Tr, Hom, Sym, As mean is the arithmetic mean A, which is Un, Int, and

[6] Der Satz vom arithmetischen Mittel in axiomatischer Begründung, Math. Ann., vol. 68 (1910), pp. 125–132.

Cont. Indeed, Schimmack proved (p. 129) that a Tr, Hom ($k = -1$), Sym mean of *two* elements is A.

For Mon means—and hence for Inc means—the following simple theorem will be proven:

THEOREM IV. *A unique monotone mean $m = F(x_1, x_2, \cdots)$ is internal, provided that for each real x,*

$$F(x, x, \cdots) = x. \tag{2}$$

Proof. Suppose, if possible, that the single-valued mean m is external; say

$$m = F(x_1, x_2, \cdots) = \mu + c,$$

where $\mu = \text{Max}\,(x_1, x_2, \cdots)$; $c > 0$. Then, since m is monotone,

$$F(\mu, \mu, \cdots) \geqq F(x_1, x_2, \cdots) = \mu + c.$$

But, by hypothesis, $F(\mu, \mu, \cdots) = \mu$,—in contradiction with the foregoing inequality. Likewise, the assumption that $m < \text{Min}\,(x_1, x_2, \cdots)$ leads to a contradiction.

COROLLARY. *A mean monotone on a branch is internal on that branch if mean $(x, x, \cdots, x) = x$.*

§4. Illustrations

1. Any student of the mathematics of finance is familiar with the equated time m for discharging debts $c_1, c_2, \cdots, c_n$, due in $x_1, x_2, \cdots, x_n$ years, respectively. The mean m is found from

$$c(1 + i)^{-m} = \sum c_r(1 + i)^{-x_r}; \qquad c_r > 0, \qquad c = \sum c_r, \tag{5}$$

where i is the annual (effective) rate of interest. This is—see (4)—a weighted exponential mean (E') with the base $b = (1 + i)^{-1}$; it is increasing (Inc) and translative (Tr), and it is associative (As) if the c's are all equal; but it is not homogeneous (Hom). If we wish to take a half-year as a unit of time, we may replace the foregoing equation by

$$c(1 + j/2)^{-2m} = \sum c_r(1 + j/2)^{-2x_r}, \tag{6}$$

where j is the corresponding *nominal* interest rate, payable semi-annually. Then, the $2m$ half-years in (6) is the same interval of time as the m years in (5). But, simply to double each x_r in (5) does not double m; this mean is *not homogeneous*—and nevertheless useful.

2. Another important mean in finance is the composite life of an industrial plant. For a machine of wearing value c_r and expected life x_r, a sinking fund deposit $c_r/s_{\overline{x_r}|}$ is made. The composite life is then the mean m of expected lives x_r, where

$$c/s_{\overline{m}|} = \sum c_r/s_{\overline{x_r}|}; \qquad is_{\overline{m}|} = (1 + i)^m - 1. \tag{7}$$

This mean is Inc; but it is neither Hom nor Tr. It is As if the c_r's are all equal.

3. From n logs of equal length and equal diameter d, n beams have been cut with widths $x_1, x_2, \cdots x_n$. What is the mean width of the beams, relative to strength—assuming that the strength of a beam varies as its width and the square of its depth? By hypothesis, the strength of the beam of width x_i is

$$\psi(x_i) = K[x_i d^2 - x_i^3] ; \qquad K = \text{constant}. \tag{8}$$

We are to find a mean m satisfying (3), which may be written

$$\sum \psi(x_i) = n\psi(m) . \tag{3'}$$

Note that ψ is constructed with reference to the *use* that is made of the mean. Now in the interval from 0 to d, $y = \psi(x)$ takes its maximum when $x = d/\sqrt{3}$. In case each $x = d/\sqrt{3}$, $m = d/\sqrt{3}$. Otherwise, there exist exactly two means m and m', with $m < d/\sqrt{3}$, and $m' > d/\sqrt{3}$. The case where each $x_i < d/\sqrt{3}$ is of special interest; since then m' is *external*. That is, we may cut n beams with width m', *greater than every given width* x_i; and the combined strength of the new set of n beams will be the same as that of the given set.

4. The spoke of a wheel has revolved from a horizontal position through an angle x_i. What is the mean m of n of such angles, relative to the slope of the spoke in the final position—supposing no spoke vertical? Find m so that

$$n \tan m = \sum \tan x_i . \tag{9}$$

The mean m is *infinitely many-valued*. On each branch it is Inc. Moreover, it is As. Reference to the graphical significance of the As mean at the close of §1 will be helpful. We may select any r of the n angles x_i; and for these r angles, an infinitely multiplied-valued mean m' exists. In place of each of the r selected angles, we may now substitute one of the values of m', a different one in each case if we like. Such substitution leaves unaltered the original many-valued mean m.

5. To blocks of ice, all of the same weight and at 0°F, heat is applied—in amount $x_1, x_2, \cdots, x_n$ units, respectively. What is the mean amount applied, relative to change in temperature? Let the temperatures attained be $y_1, y_2, \cdots, y_n$, degrees respectively. We require the amount of heat m, which applied to each block of ice at 0° would give it the temperature $\sum y_i/n$. If this temperature should be 32°—for freezing—the m may have any v alue over a certain finite range, on account of latent heat. Likewise, for the boiling temperature, 212°. Such values of x are points of *discontinuity* for m.

On the other hand, we may seek the mean final temperature, relative to quantities of heat absorbed. We then require the temperature m' to be attained if each block of ice is to absorb $\sum x_i/n$ units of heat. This mean m' likewise has points of discontinuity.

6. Let the sum of money c_r be due in r years; and let it be offered for

$$c'_r = c_r(1 + i_r)^{-r} . \tag{10}$$

What mean interest rate i will be realized if several of these obligations c_r are purchased? Find i so that—with summation now with respect to r—

$$\sum c_r(1+i)^{-r} = \sum c'_r = \sum c_r(1+i_r)^{-r}. \tag{11}$$

This mean i of given interest rates i_r is Un, Int, and Inc.

This kind of mean—more general than (3) in that n different functions $\psi_r(x_r) = (1+i_r)^{-r}$ are used, and with weights c_r—I defined and discussed in a former paper.[7] If, now, the obligations are made infinite in number, a species of *perpetuity* is formed. Under rather simple conditions, such a series would converge, and a unique mean interest rate i would be determined, a mean of a *countable* set of interest rates, i_r.

For further illustrations of means, the reader is referred to the papers of Chisini and De Finetti already cited, and to papers of P. Martinotti[8] and C. Gini-L. Galvani.[9]

§5. Summary

The mean $m = F(x_1, x_2, \cdots)$ is required to satisfy but a single condition, namely: The function F must be such that $F(x, x, \cdots) = x$.

Such a mean may be infinitely-many valued, it may be external to all given values $x_1, x_2, \cdots, x_n$, it may be discontinuous.

Among the properties designated by internal, unique, homogeneous, translative, symmetric, increasing, monotone (increasing), associative, and continuous, some dependence has already been discovered; and to this a slight addition is made in §3. But (§2) the first five of the foregoing properties are completely independent; also the set from the second to the sixth, inclusive; and some other sets.

Austin, Texas.

[7] Functions of measurements under general laws of error, Skandinavisk Aktuarietidskrift, vol. 5 (1922), pp. 133–158.

[8] Le medie relative, Giornale degli Economisti e Rivista di Statistica, ser. 4, vol. 71 (1931), pp. 291–303.

[9] Di talune estensioni dei concetti di media ai caratteri qualitativi, Metron, vol. 8 (1929), pp. 5–209.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

APPLICATIONS OF AN ABSTRACT EXISTENCE THEOREM TO BOTH DIFFERENTIAL AND DIFFERENCE EQUATIONS

By O. K. Bower

(Received March 27, 1933)

A functional equation of the form $f = g + Sf$, in which g is a known function and S a suitable functional operator, has for some time been used to prove existence theorems for integral and differential equations.[1] It is the purpose of this paper to consider an equation of this form with modified restrictions on the function g and operator S and prove an abstract existence theorem which can be applied directly to proving the existence of solutions of integral, finite integral, q-integral, differential, difference, and q-difference equations having relatively simple solutions in the neighborhood of infinity.

1. Abstract Existence Theorem. The general region R of which several particular cases will be considered is the entire finite portion of the complex plane outside some combination of circles, segments, and lines.

Definition: *A function $\phi(z)$ is of class K in a region R of the complex plane if $\phi(z)$ is analytic in R and the limit of $z^k\phi(z)$ as z approaches infinity in R exists.*

Let S be a linear[2] operator which satisfies the following postulates:

(a) If $\phi(z)$ is of class K in R_1 then $S\phi$ exists and is of class K in R_1.

(b) If $\phi(z)$ is of class K in R_1 there exists in R_1 a region R_2, the same for all ϕ, such that the infinite series $\phi + S\phi + S^2\phi + \cdots$ converges for every value of z in R_2 and represents a function of class K in R_2.

(c) If the sum of the series in (b) is $\psi(z)$ then $S\psi = S\phi + S^2\phi + S^3\phi + \cdots$.

Theorem. *The equation*

$$f = g + Sf, \tag{1}$$

in which $g(z)$ is of class K in a region R_1 and S satisfies the above postulates with respect to functions of class K in R_1 has a unique solution of class K in the region R_2 contained in R_1.

For suppose equation (1) has a solution $f(z)$ of class K in R_2. Then by postulate (b) $\lim_{n=\infty} S^n f = 0$, and by repeated substitutions of $g + Sf$ for f in the second member of (1) it is clear that $f(z)$ may be written $f(z) = g + Sg + S^2g + \cdots$. Furthermore there exists a function $f(z) = g + Sg + S^2g + \cdots$ of class K in R_2, and this satisfies equation (1), since $Sf = Sg + S^2g + \cdots = f - g$.

[1] See Max Mason, "The New Haven Mathematical Colloquium," Yale University Press, 1910.

[2] An operator T is said to be linear if, when Tu and Tv exist, $T(u + v) = Tu + Tv$ and $Tcu = cTu$, c being a constant.

2. **Application of Abstract Theorem to an Integral Equation.** Let the region R_1 be the entire finite plane outside a circle of radius a about the point $z = 0$ as center, or outside the circle and a sector with vertex at zero. Let $g(z)$ be of class 2 in R_1, and let $P(z, t)$ be a function analytic in z and t simultaneously for z in R_1 and admissible values of t, that is, for values of t in R_1 such that $|t| \geqq |z|$, and such that there exists a positive constant N for which $|P(z, t)| \leqq N|z|^{-2}$ uniformly for z in R_1 and admissible values of t.

Consider the integral equation

$$f(z) = g(z) + \int_z^\infty P(z, t) f(t)\, dt \equiv g + S_1 f, \tag{2}$$

z being any point in R_1, and the path of integration being along the straight line through the points 0, z. *Equation* (2) *has a unique solution of class* 2 *in* R_1*:*

$$f(z) = g(z) + S_1 g(z) + \cdots = g + \int_z^\infty P(z, t) g(t)\, dt + \cdots. \tag{3}$$

Proof: If $\phi(z)$ is of class 2 in R_1 there exists a positive constant M such that for z in R_1 $|\phi(z)| \leqq M|z|^{-2}$. Then $|S_1^k \phi| \leqq N^k M/k!\, |z|^{k+2}$, and $\sum_k S_1^k \phi$ is absolutely and uniformly convergent in R_1, representing a function $\psi(z)$ of class 2 in R_1. If $\psi(z)$ is written $\phi + S_1\phi + S_1^2\phi + \cdots$ it may easily be shown that $S_1\psi = S_1\phi + S_1^2\phi + \cdots$.

3. **Proof of Existence of Solution of a Differential Equation.** Definition of an operator θ:

$$\theta^{(0)} y = y,$$

$$\theta^{(k)} y = z^2 D_z \theta^{(k-1)} y$$

$$= z^2\left[z^{2k-2} y^{(k)} + (k-1)kz^{2k-3}y^{k-1} + \frac{(k-2)(k-1)^2 k}{2!} z^{2k-4} y^{(k-2)} + \cdots + \frac{2\cdot 3^2\cdot 4^2 \cdots (k-1)^2 k}{(k-2)!} z^k y'' + k!\, z^{k-1} y'\right],$$

$$k = 1, 2, \cdots, n-1.$$

Consider the equation

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \cdots + p_{n-1} y' + p_n y = e(z) \tag{4}$$

with initial conditions

$$\lim_{z=\infty} \theta^{(k)} y = (-1)^k k!\, \alpha_k, \qquad k = 0, 1, 2, \cdots, n-1. \tag{5}$$

If both members of equation (4) are multiplied by z^{2n-2} the resulting equation can be written

$$D_z\theta^{(n-1)}y + \sum_{i=1}^{n} P_i\theta^{(n-i)}y = E(z)\,, \tag{6}$$

in which P_i can be expressed linearly in terms of $p_1, \cdots, p_n$, and

$$E = z^{2n-2}e(z)\,.$$

The conditions on the coefficients in (4) are that P_i, E are of class 2 in a region R_1 which is the part of the finite plane outside a circle of radius a about the point $z = 0$ as center. Excluding the identically zero solution, if it exists, suppose the system (4) and (5) has a solution analytic in R_1. Then this is also a solution of the system (6) and (5), and $D_z\theta^{(n-1)}y$ exists and is of class 2 in R_1. Let

$$v(z) = D_z\theta^{(n-1)}y = E - \sum_{i=1}^{n} P_i\theta^{(n-i)}y\,. \tag{7}$$

If (7) and the initial conditions (5) are used the following relations are obtained:

$$\begin{aligned}\theta^{(n-i)}y = (-1)^{n-i}\Big[(n-i)!\,\alpha_{n-i} &+ \frac{(n-i+1)!}{z}\alpha_{n-i+1} + \frac{(n-i+2)!}{2!\,z^2}\alpha_{n-i+2} \\ &+ \cdots + \frac{(n-1)\,!}{(i-1)!\,z^{i-1}}\alpha_{n-1}\Big] + \frac{(-1)^i}{(i-1)!}\int_z^\infty \left(\frac{1}{z}-\frac{1}{t}\right)^{i-1} v(t)\,dt\,, \\ &\qquad i = 1, 2, \cdots, n\,,\end{aligned} \tag{8}$$

the path of integration lying in the region R_1. When these expressions are substituted in (7) an equation is obtained in the form

$$v(z) = v_1(z) + \int_z^\infty P(z, t)\,v(t)\,dt\,, \tag{9}$$

in which $v_1(z)$ and $P(z, t)$ are such that this equation is of the type (2), and therefore has a unique solution of class 2 in R_1. The solution of (9) substituted in (8) gives the unique solution y of class 2 in R_1 for the system (4) and (5).

4. **Application of Abstract Theorem to a Finite Integral Equation. Solution to the Right.** Let R_1 be the part of the finite plane outside a circle of radius a and two parallel half lines extending to the left from the points ai and $-ai$ inclosing the left half of the axis of reals. The "symbolic solution to the right" of the equation $\Delta y = \phi(z)$ is $-\sum_{k=0}^{\infty}\phi(z+k)$, which will be written $-\mathcal{S}_z^\infty \phi(t)$.

Consider the finite integral equation

$$f(z) = g(z) + \mathcal{S}_z^\infty P(z, t)\,f(t) \equiv g(z) + S_2 f\,, \tag{10}$$

in which $g(z)$ is of class 2 in R_1, and $P(z, t)$ is analytic in R_1 simultaneously with respect to z and admissible values of t, that is, for values of t in R_1 such that $|t| \geqq |z|$, and such that there exists a positive constant N for which $|P(z, t)| \leqq N|z|^{-2}$ uniformly for z in R_1 and admissible values of t. There is a region R_2 contained in R_1 and with boundary equidistant from the boundary of R_1 in which (10) has a unique solution of class 2.

PROOF: As for equation (3), $|\phi(z)| \leqq M|z|^{-2}$ for z in R_1. In the same region

$$|S_2\phi| \leqq N/|z|^2 \sum_{k=0}^{\infty} |\phi(z+k)| \leqq N/|z|^2 \sum_{k=0}^{\infty} M/|z+k|^2.$$

If the boundary of R_2 is chosen sufficiently far from the boundary of R_1 and the part of R_2 in the right half plane is denoted by R_2^+ and the part in the left half plane by R_2^-, by a theorem in difference equations[3] there exists a positive constant C such that if $m \geqq 2$, $z = u + iv$,

$$\textstyle\sum_{k=0}^{\infty} |z+k|^{-m} < C/|z|^{m-1}, \qquad \sum_{k=0}^{\infty} |z+k|^{-m} < C/|v|^{m-1}$$

for z in R_2^+ and R_2^- respectively. Then

$$|S_2^k\phi| < N^k MC^k/|z|^{k+2},$$

$$|S_2^k\phi| < N^k MC^k/|z|^2 v^k$$

for z in R_2^+ and R_2^- respectively. It follows that $\Sigma_k S_2^k\phi$ is absolutely and uniformly convergent in R_2 and represents a function

$$\psi(z) = \phi + S_2\phi + S_2^2\phi + \cdots.$$

Furthermore

$$S_2\psi = S_2\phi + S_2^2\phi + \cdots.$$

5. **Application of Abstract Theorem to a Finite Integral Equation. Solution to the Left.** The region $\bar{R}_1$ is the part of the finite plane outside a circle of radius a about the point $z = 0$ as center and two parallel half lines extending to the right from the points ai, $-ai$, inclosing the right half of the axis of reals. The region $\bar{R}_2$ is contained in $\bar{R}_1$ with its boundary equidistant from the boundary of $\bar{R}_1$ at a distance sufficiently great. The "symbolic solution to the left" of the equation $\Delta y = \phi(z)$ is $\sum_{k=1}^{\infty} \phi(z-k)$, and is written $\mathop{\mathcal{S}}_{z-1}^{-\infty} \phi(t)$.

In a manner similar to that above the finite integral equation

$$f(z) = g(z) + \mathop{\mathcal{S}}_{z-1}^{-\infty} P(z, t)\, f(t) \equiv g(z) + S_3 f, \tag{11}$$

[3] See, for instance, P. M. Batchelder, "Introduction to Linear Difference Equations," 1927, p. 25.

If both members of equation (4) are multiplied by z^{2n-2} the resulting equation can be written

$$D_z\theta^{(n-1)}y + \sum_{i=1}^{n} P_i\theta^{(n-i)}y = E(z)\,, \tag{6}$$

in which P_i can be expressed linearly in terms of $p_1, \cdots, p_n$, and

$$E = z^{2n-2}e(z)\,.$$

The conditions on the coefficients in (4) are that P_i, E are of class 2 in a region R_1 which is the part of the finite plane outside a circle of radius a about the point $z = 0$ as center. Excluding the identically zero solution, if it exists, suppose the system (4) and (5) has a solution analytic in R_1. Then this is also a solution of the system (6) and (5), and $D_z\theta^{(n-1)}y$ exists and is of class 2 in R_1. Let

$$v(z) = D_z\theta^{(n-1)}y = E - \sum_{i=1}^{n} P_i\theta^{(n-i)}y\,. \tag{7}$$

If (7) and the initial conditions (5) are used the following relations are obtained:

$$\begin{aligned}\theta^{(n-i)}y = (-1)^{n-i}\Big[(n-i)!\,\alpha_{n-i} &+ \frac{(n-i+1)!}{z}\alpha_{n-i+1} + \frac{(n-i+2)!}{2!\,z^2}\alpha_{n-i+2} \\ &+ \cdots + \frac{(n-1)!}{(i-1)!\,z^{i-1}}\alpha_{n-1}\Big] + \frac{(-1)^i}{(i-1)!}\int_z^\infty \left(\frac{1}{z}-\frac{1}{t}\right)^{i-1} v(t)\,dt\,,\end{aligned} \tag{8}$$

$$i = 1, 2, \cdots, n\,,$$

the path of integration lying in the region R_1. When these expressions are substituted in (7) an equation is obtained in the form

$$v(z) = v_1(z) + \int_z^\infty P(z,t)\,v(t)\,dt\,, \tag{9}$$

in which $v_1(z)$ and $P(z, t)$ are such that this equation is of the type (2), and therefore has a unique solution of class 2 in R_1. The solution of (9) substituted in (8) gives the unique solution y of class 2 in R_1 for the system (4) and (5).

4. **Application of Abstract Theorem to a Finite Integral Equation. Solution to the Right.** Let R_1 be the part of the finite plane outside a circle of radius a and two parallel half lines extending to the left from the points ai and $-ai$ inclosing the left half of the axis of reals. The "symbolic solution to the right" of the equation $\Delta y = \phi(z)$ is $-\sum_{k=0}^{\infty} \phi(z+k)$, which will be written $-\mathcal{S}_z^\infty \phi(t)$.

Consider the finite integral equation

$$f(z) = g(z) + \mathcal{S}_z^\infty P(z,t)\,f(t) \equiv g(z) + S_2 f\,, \tag{10}$$

in which $g(z)$ is of class 2 in R_1, and $P(z, t)$ is analytic in R_1 simultaneously with respect to z and admissible values of t, that is, for values of t in R_1 such that $|t| \geqq |z|$, and such that there exists a positive constant N for which $|P(z,t)| \leqq N|z|^{-2}$ uniformly for z in R_1 and admissible values of t. There is a region R_2 contained in R_1 and with boundary equidistant from the boundary of R_1 in which (10) has a unique solution of class 2.

PROOF: As for equation (3), $|\phi(z)| \leqq M|z|^{-2}$ for z in R_1. In the same region

$$|S_2\phi| \leqq N/|z|^2 \sum_{k=0}^{\infty} |\phi(z+k)| \leqq N/|z|^2 \sum_{k=0}^{\infty} M/|z+k|^2.$$

If the boundary of R_2 is chosen sufficiently far from the boundary of R_1 and the part of R_2 in the right half plane is denoted by R_2^+ and the part in the left half plane by R_2^-, by a theorem in difference equations[3] there exists a positive constant C such that if $m \geqq 2$, $z = u + iv$,

$$\textstyle\sum_{k=0}^{\infty} |z+k|^{-m} < C/|z|^{m-1}, \qquad \sum_{k=0}^{\infty} |z+k|^{-m} < C/|v|^{m-1}$$

for z in R_2^+ and R_2^- respectively. Then

$$|S_2^k\phi| < N^kMC^k/|z|^{k+2},$$

$$|S_2^k\phi| < N^kMC^k/|z|^2 v^k$$

for z in R_2^+ and R_2^- respectively. It follows that $\Sigma_k S_2^k\phi$ is absolutely and uniformly convergent in R_2 and represents a function

$$\psi(z) = \phi + S_2\phi + S_2^2\phi + \cdots.$$

Furthermore

$$S_2\psi = S_2\phi + S_2^2\phi + \cdots.$$

5. **Application of Abstract Theorem to a Finite Integral Equation. Solution to the Left.** The region $\bar{R}_1$ is the part of the finite plane outside a circle of radius a about the point $z = 0$ as center and two parallel half lines extending to the right from the points ai, $-ai$, inclosing the right half of the axis of reals. The region $\bar{R}_2$ is contained in $\bar{R}_1$ with its boundary equidistant from the boundary of $\bar{R}_1$ at a distance sufficiently great. The "symbolic solution to the left" of the equation $\Delta y = \phi(z)$ is $\sum_{k=1}^{\infty} \phi(z-k)$, and is written $\mathcal{S}_{z-1}^{-\infty} \phi(t)$.

In a manner similar to that above the finite integral equation

$$f(z) = g(z) + \mathcal{S}_{z-1}^{-\infty} P(z,t)\, f(t) \equiv g(z) + S_3 f, \tag{11}$$

[3] See, for instance, P. M. Batchelder, "Introduction to Linear Difference Equations," 1927, p. 25.

in which $g(z)$ is of class 2 in $\bar{R}_1$, and $P(z, t)$ is analytic in $\bar{R}_1$ simultaneously with respect to z and admissible values of t, that is, for values of t in $\bar{R}_1$ such that $|t| \geqq |z|$, and such that there exists a positive constant N for which $|P(z,t)| \leqq N|z|^{-2}$ uniformly for z in $\bar{R}_1$ and admissible values of t, may be shown to possess a unique solution of class 2 in the region $\bar{R}_2$.

6. Proof of Existence of Solution of a Difference Equation in the Region R_2. Definition of an operator Γ_n:

$$\Gamma_n^{(0)} y = y ,$$

$$\Gamma_n^{(k)} y = (z + n - k - 1)(z + n - k)\, \Delta\, \Gamma_n^{(k-1)} y$$

$$= (z + n - k - 1)(z + n - k)\, [(z + n - k + 1) \cdots (z + n + k - 2)\, \Delta^k y$$

$$+ \cdots + k!\,(z + n - k + 1) \cdots (z + n - 1)\, \Delta y] ,$$

$$k = 1, 2, \cdots, n - 1 .$$

Consider the equation

$$\Delta^n y + p_1 \Delta^{n-1} y + \cdots + p_{n-1} \Delta y + p_n y = e(z) \tag{12}$$

with initial conditions, the limit to be taken as z approaches infinity in a region R_2 of the type in section 4 above,

$$\lim \Gamma_n^{(k)} y = (-1)^k k!\, \alpha_k , \qquad k = 0, 1, 2, \cdots, n - 1 . \tag{13}$$

If both members of equation (12) are multiplied by

$$(z + 1)(z + 2) \cdots (z + 2n - 2)$$

the resulting equation can be written

$$\Delta\, \Gamma_n^{(n-1)} y + \sum_{i=1}^{n} P_i\, \Gamma_n^{(n-i)} y = E(z) , \tag{14}$$

in which P_i can be expressed linearly in terms of

$$p_1, \cdots, p_n , \quad \text{and} \quad E = (z + 1)(z + 2) \cdots (z + 2n - 2)\, e(z) .$$

The conditions on the coefficients in (12) are that P_i and E are of class 2 in a region R_1 of the type in section 4. Excluding the identically zero solution, if such exists, suppose that (12) has a solution analytic in R_2 and this solution satisfies the conditions (13). Then this solution is also a solution of the system (14) and (13), and $\Delta\, \Gamma_n^{(n-1)} y$ exists and is of class 2 in R_2. Let

$$w(z) = \Delta\, \Gamma_n^{(n-1)} y = E - \sum_{i=1}^{n} P_i\, \Gamma_n^{(n-i)} y . \tag{15}$$

If (15) and the initial conditions (13) are used the following relations are obtained:

$$\Delta\, \Gamma_n^{(n-1)}\, y = w(z)$$

$$\Gamma_n^{(n-i)}\, y = (-1)^{n-1}\,(n-1)\,!\,\alpha_{n-1} - \mathop{\mathbf{S}}_z^{\infty} w(t)$$

$$\begin{aligned}
\Gamma_n^{(n-i)}\, y = (-1)^{n-i}\Bigg[&(n-i)\,!\,\alpha_{n-i} + \frac{(n-i+1)\,!}{z+i-2}\,\alpha_{n-i+1} \\
&+ \frac{(n-i+2)\,!}{2\,!\,(z+i-3)\,(z+i-2)}\,\alpha_{n-i+2} + \cdots \\
&+ \frac{(n-1)\,!}{(i-1)\,!\,z(z+1)\cdots(z+i-2)}\,\alpha_{n-1}\Bigg] \\
&+ \frac{(-1)^i}{(i-1)\,!}\mathop{\mathbf{S}}_z^{\infty}\Bigg[\frac{1}{z(z+1)\cdots(z+i-2)} \\
&- \frac{i-1}{(z+1)\cdots(z+i-2)\,(t+1)} \\
&+ \frac{\binom{i-1}{2}}{(z+2)\cdots(z+i-2)\,(t+1)\,(t+2)} - \cdots \\
&+ \frac{(-1)^{i-1}}{(t+1)\cdots(t+i-1)}\Bigg]\,w(t)\,, \\
&\qquad\qquad i = 2\,,\cdots,n\,.
\end{aligned} \tag{16}$$

When these expressions for $\Gamma_n^{(n-i)}\, y$ are substituted in (15) an equation is obtained in the form

$$w(z) = w_1(z) + \mathop{\mathbf{S}}_z^{\infty} P(z,t)\, w(t)\,, \tag{17}$$

in which $w_1(z)$ and $P(z, t)$ are such that this equation is of type (10), and therefore has a unique solution of class 2 in the region R_2. The solution of (17) substituted in (16) gives the unique solution y of the system (12) and (13).

7. Proof of Existence of Solution of a Difference Equation in the Region $\overline{R_2}$. Definition of an operator $\bar{\Gamma}_n$:

$$\bar{\Gamma}_n^{(0)}\, y = y$$

$$\bar{\Gamma}_n^{(1)}\, y = (z-n)\,(z-n+1)\,\Delta\, y$$

$$\begin{aligned}
\bar{\Gamma}_n^{(k)}\, y &= (z-n-k+1)\,(z-n-k+2)\,\Delta\,\bar{\Gamma}_n^{(k-1)}\, y \\
&= (z-n-k+1)\,(z-n-k+2)\,[(z-n-k+3)\cdots(z-n+k)\,\Delta^k y \\
&\qquad + \cdots + k\,!\,(z-n-k+3)\cdots(z-n+1)\,\Delta\, y]\,, \\
&\qquad\qquad k = 2\,,\cdots,n-1\,.
\end{aligned}$$

Consider the equation

$$\Delta^n y + p_1 \Delta^{n-1} y + \cdots + p_{n-1} \Delta y + p_n y = e(z) \tag{18}$$

with initial conditions, the limit to be taken as z approaches infinity in a region $\bar{R}_2$ of the type in section 5 above:

$$\lim \bar{\Gamma}_n^{(k)} y = (-1)^k k!\, \alpha_k, \quad k = 0, 1, \cdots, n-1. \tag{19}$$

If both members of equation (18) are multiplied by

$$(z - 2n + 3)(z - 2n + 4) \cdots (z - 1)z$$

the resulting equation can be written

$$\Delta \bar{\Gamma}_i^{(n-1)} y + \sum_{i=1}^{n} P_i \bar{\Gamma}_n^{(n-i)} y = E(z), \tag{20}$$

in which P_i can be expressed linearly in terms of

$$p_1, \cdots, p_n, \quad \text{and} \quad E = (z - 2n + 3)(z - 2n + 4) \cdots (z - 1) z\, e(z).$$

The conditions on the coefficients in (18) are that P_i and E are of class 2 in a region $\bar{R}_1$ of the type in section 5. Excluding the identically zero solution, if such exists, suppose that (18) has a solution analytic in $\bar{R}_2$, a region of the type in section 5, and this solution satisfies the conditions (19). Then this solution is also a solution of the system (20) and (19), and $\Delta \bar{\Gamma}_n^{(n-1)} y$ exists and is of class 2 in $\bar{R}_2$. Let

$$\bar{w}(z) = \Delta \bar{\Gamma}_n^{(n-1)} y = E - \sum_{i=1}^{n} P_i \bar{\Gamma}_n^{(n-i)} y. \tag{21}$$

If (22) and the initial conditions (20) are used the following relations are obtained:

$$\Delta \bar{\Gamma}_n^{(n-1)} y = \bar{w}(z)$$

$$\bar{\Gamma}_n^{(n-1)} y = (-1)^{n-1} (n-1)!\, \alpha_{n-1} + S_{z-1}^{-\infty} \bar{w}(t)$$

$$\begin{aligned}\bar{\Gamma}_n^{(n-i)} y = (-1)^{n-i} \Bigg[& (n-i)!\, \alpha_{n-i} + \frac{(n-i+1)!}{z-2n+i} \alpha_{n-i+1} \\ & + \frac{(n-i+2)!\ \alpha_{n-i+2}}{2!\,(z-2n+i-1)(z-2n+i)} \\ & + \cdots + \frac{(n-1)!\, \alpha_{n-1}}{(i-1)!\,(z-2n+2)\cdots(z-2n+i)} \Bigg] \\ + \frac{1}{(i-1)!} S_{z-1}^{-\infty} \Bigg[& \frac{1}{(t-2n+3)\cdots(t-2n+i+1)} \\ & - \frac{i-1}{(t-2n+3)\cdots(t-2n+i)(z-2n+i)} \\ & + \cdots + \frac{(-1)^{i-1}}{(z-2n+2)\cdots(z-2n+i)} \Bigg] \bar{w}(t), \quad i = 2, \cdots, n.\end{aligned} \tag{22}$$

When these expressions for $\bar{\Gamma}_n^{(n-i)} y$ are substituted in (21) an equation is obtained in the form

$$\bar{w}(z) = \bar{w}_1(z) + \mathop{\mathcal{S}}_{z-1}^{-\infty} P(z, t)\, \bar{w}(t)\,, \tag{23}$$

in which $\bar{w}_1(z)$ and $P(z, t)$ are such that this equation is of type (11) and therefore has a unique solution of class 2 in the region R_2. The solution of (23) substituted in (22) gives the unique solution y of class 2 of the system (18) and (19).

8. **Analytic Continuation and Asymptotic Expansions.** The equations (12) and (18) may be written

$$\begin{aligned} - 1/A_n\, y(z + n) + A_1/A_n\, y(z + n - 1) + \cdots + A_{n-1}/A_n\, y(z + 1) \\ + e(z)/A_n = y(z)\,, \end{aligned} \tag{12a}$$

and

$$y(z) = B_1\, y(z - 1) + B_2\, y(z - 2) + \cdots + b_n\, y(z - n) + e(z - n) \tag{18a}$$

respectively. By means of these equations the solutions y may be continued analytically into any part of the plane, and the extended solutions will be analytic except possibly in (12a) for singularities and points congruent to them on the left of the functions $-1/A_n$, A_i/A_n, $e(z)/A_n$, and except possibly in (18a) for singularities and points congruent to them on the right of the functions B_i, $e(z - n)$.

When the coefficients in equation (12) are analytic in the neighborhood of infinity it can be shown that the solution y has an asymptotic expansion in that part of the complex plane outside a small sector containing the axis of reals and lying in R_2. This asymptotic expansion is the formal solution in descending power series in equation (12) using the initial conditions (13). It can also be shown that the solution of (18) and (19) has an asymptotic expansion in a suitable sector which excludes the positive axis of reals.

9. **Application of Abstract Theorem to a q-Integral Equation.** Let R_1 be the part of the finite plane which is outside a circle of radius a about the point $z = 0$ as center. The equation $\delta f(z) = f(qz) - f(z) = \phi(z)$ has for one of its symbolic solutions $- \sum_{i=0}^{\infty} \phi(q^i z)$, which will be written $- \mathop{\mathcal{S}}_{z}^{\infty}{}_{q}\, \phi(t)$. Let $|g| > 1$, and consider the q-integral equation

$$f(z) = g(z) + \mathop{\mathcal{S}}_{z}^{\infty}{}_{q}\, P(z, t)\, f(t) \equiv g(z) + S_4 f, \tag{24}$$

in which $g(z)$ is of class 1 in R_1, and $P(z, t)$ is analytic simultaneously with respect to z in R_1 and admissible values of t, that is, for values of t in R_1 such that $|t| \geqq |z|$, and such that there exists a positive constant N for which

$| P(z, t) | \leqq N | z |^{-1}$ uniformly for z in R_1 and admissible values of t. *The equation* (24) *has a unique solution of class* 1 *in a region* R_2 *which is all of* R_1 *outside a circle of radius greater than* N *about* $z = 0$ *as center.*

PROOF: If $\phi(z)$ is of class 1 in R_1 there exists a positive constant M such that $| \phi(z) | \leqq M | z |^{-1}$ for z in R_1. Then the following inequalities exist for $| S_4^k \phi |$:

$$| S_4^k \phi | \leqq N^k M / | z |^{k+1} \cdot \frac{| q |}{| q | - 1} \cdot \frac{| q^2 |}{| q^2 | - 1} \cdots \frac{| q^k |}{| q^k | - 1}$$

for $| z | > N$. Hence in the region R_2 the series

$$\phi + S_4 \phi + S_4^2 \phi + \cdots$$

converges absolutely and uniformly, and represents a function

$$\psi(z) = \phi + S_4 \phi + S_4^2 \phi + \cdots$$

of class 1 in R_2. Furthermore

$$S_4 \psi = S_4 \phi + S_4^2 \phi + \cdots .$$

10. **Proof of Existence of a Solution of a q-Difference Equation.** In the following definition, for brevity of notation, the symbol

$$\sum_1^r \frac{q^{s1} - 1}{q^{s1}} \cdot \frac{q^{s2} - 1}{q^{s2}} \cdots \frac{q^{sk} - 1}{q^{sk}}$$

will be written C_{rk}, and represents the symmetric polynomial in the quantities $q^{-1}/q, q^2 - 1/q^2, \cdots, q^r - 1/q^r$ which consists of the sum of the products of these r quantities taken k at a time. Definition of an operator Γ_q:

$$\Gamma_q^{(0)} y = y$$

$$\Gamma_q^{(i)} y = qz/q^{i-1} \cdot \delta \Gamma_q^{(i-1)} y = q^i/q^i - 1 \cdot q^{i-1}/q^{i-1} - 1 \cdots q/q - 1 \cdot q^{\frac{1}{2}i(i-1)} z^i [\delta^i y + C_{i-1,1} \delta^{i-1} y + C_{i-1,2} \delta^{i-2} y + \cdots + C_{i-1,i-1} \delta y], \quad i = 1, 2, \cdots, n-1.$$

Consider the equation

$$\delta^n y + p_1 \delta^{n-1} y + \cdots + p_n y = e(z) \tag{25}$$

with initial conditions

$$\lim_{z=\infty} \Gamma_q^{(k)} y = (-1)^k \alpha_k, \qquad k = 0, 1, \cdots, n-1. \tag{26}$$

If both members of equation (26) are multiplied by

$$q^{n-1}/q^{n-1} - 1 \cdots q/q - 1 \cdot q^{\frac{1}{2}n(n-1)} z^{n-1}$$

the resulting equation can be written

$$\delta \Gamma_q^{(n-1)} y + \sum_{i=1}^n P_i \Gamma_q^{(n-i)} y = E(z), \tag{27}$$

in which, $B_k(z)$ being a function which satisfies the recurrence relation

$$B_k = p_k - \sum_{i=0}^{k-1} B_i C_{n-i-1,k-i}, \qquad B_0 = 1,$$

$$P_k = q^{n-1}/q^{n-1} - 1 \cdots q^{n-k+1}/q^{n-k+1} - 1 \cdot q^{\frac{1}{2}k(2n-k-1)} z^{k-1} \{B_k(z)\},$$

$$E = q^{n-1}/q^{n-1} - 1 \cdots q/q - 1 \cdot q^{\frac{1}{2}n(n-1)} z^{n-1} e(z).$$

The conditions on the coefficients in (25) are that P_i and E are of class 1 in the region R_1 which is the finite plane outside a circle of radius a about the point $z = 0$ as center. Excluding the identically zero solution, if such exists, suppose the system (25) and (26) has a solution analytic in a region R_2 which is the finite plane outside a circle of sufficiently large radius concentric with the circle which is the boundary of R_1. Then this solution is also a solution of (27) with conditions (26), and $\delta\Gamma_q^{(n-1)} y$ exists and is of class 1 in R_2. Let

$$\pi(z) = \delta\Gamma_q^{(n-1)} y = E - \sum_{i=1}^{n} P_i \Gamma_q^{(n-i)} y. \tag{28}$$

If (28) and the initial conditions (26) are used the following relations are obtained:

$$\delta\Gamma_q^{(n-1)} y = \pi(z).$$

$$\begin{aligned}\Gamma_q^{(n-k)} y = {} & (-1)^{n-k} [\alpha_{n-k} + q^{n-k+1} - 1/q^{n-k+1} \cdot q/q - 1 \cdot \alpha_{n-k+1} z^{-1} \\ & + q^{n-k+2} - 1/q^{n-k+2} \cdot q^{n-k+1} - 1/q^{n-k+1} \cdot q/q - 1 \cdot q^2/q^2 - 1\, \alpha_{n-k+2} z^{-2} \\ & + \cdots + q^{n-1} - 1/q^{n-1} \cdots q^{k-1}/q^{k-1} - 1\, \alpha_{n-1} z^{-(k-1)}] \\ & + (-1)^k z^{-(k-1)} q^{n-1} - 1/q^{n-1} \cdots q^{n-k+1} - 1/q^{n-k+1} \cdot 1/(q-1) \cdots (q^{k-1} - 1) \\ & \mathop{S}_{z}^{\infty} (qt - z)(q^2 t - z) \cdots (q^{k-1} t - z)\, t^{-(k-1)} \pi(t), \qquad k = 1, 2, \cdots, n.\end{aligned} \tag{29}$$

When these expressions for $\Gamma_q^{(n-k)} y$ are substituted in (28) an equation is obtained in the form

$$\pi(z) = \pi_1(z) + \mathop{S_q}_{z}^{\infty} P(z, t)\, \pi(t), \tag{30}$$

in which $\pi_1(z)$ and $P(z, t)$ are such that this equation is of type (24), and therefore has a unique solution of class 1 in the region R_2. It easily follows that a q-difference equation in the form (25) with initial conditions (26) at infinity, the coefficients being such that the P_i and E are of class 1 in R_1, has a unique solution analytic in R_2.

If equation (25) is put in the form

$$-1/A_n y(q^n z) + A_1/A_n y(q^{n-1} z) + \cdots + A_{n-1}/A_n y(qz) + e(z)/A_n = y(z)$$

this solution may be continued analytically back into the circle about the point $z = 0$, and will have no singularities except possibly at the singular points, λ, of the functions $-1/A_n$, Ai/A_n, $e(z)/A_n$ and points congruent toward zero, $\lambda q^{-1}, \lambda q^{-2}, \cdots$.

For $|q| < 1$, a q-integral equation which involves the operator which is suggested by the other symbolic solution $\sum_{i=1}^{\infty} \phi(q^i z)$ of the equation $\delta f(z) = \phi(z)$ can be shown to have a unique solution of class 1 in a suitable region R_2 bounded by a circle about $z = 0$ as center, and by means of this the q-difference equation can be shown to have a unique solution analytic in R_2.

UNIVERSITY OF ILLINOIS.

ANNALS OF MATHEMATICS
Vol. 35, No. 4, October, 1934

EXPANSIONS OF TWO ARBITRARY ANALYTIC FUNCTIONS IN A SERIES OF RATIONAL FUNCTIONS

BY P. W. KETCHUM

(Received, June 16, 1933)

1. It is well known that series of rational functions may converge in distinct regions, and to different analytic functions in the different regions. Weierstrass[1] and others have shown how to construct such series when the functions are arbitrarily chosen from a large class of analytic functions. Montel,[2] Borel[3] and others have shown the existence of series of rational functions, in fact, polynomials, which converge to any number of arbitrary analytic functions in distinct regions. None of these methods, however, leads to expansions in the sense of Maclaurin's expansion, since the form of the terms in the series, as well as the coefficients, depends on the function being expanded. Also the coefficients are not uniquely determined by the function being expanded.

It is the object of the present paper to develop explicit expansions analogous to Maclaurin's expansion. We study the series

$$\alpha_0 F_0(x) + \beta_0\, x\, F_0(x) + \alpha_1 F_1(x) + \beta_1\, x\, F_1(x) + \cdots, \tag{1}$$

where the functions $F_n(x)$ have in §§2, 3, 4, 5 the particular form

$$F_n(x) = x^n/(1 - x^{2n+1}), \tag{1'}$$

but are not so restricted in §6. The convergence character of the special series (1), (1′) is closely analogous to that of a Maclaurin series except that it converges in two regions. *Mere convergence at a single point, other than zero or infinity, implies convergence in two circular regions, and moreover to two analytic functions in the two regions.* More exactly, in §2 we shall prove

THEOREM I: *The series* (1), (1′) *converges absolutely and uniformly for* $|x| \leqq R'$ *and for* $|x| \geqq 1/R'$, *where* R' *is any positive number less than* R, *and* R *in turn is a number less than or equal to* 1, *which depends on the coefficients, and which may properly be called the radius of convergence of the series.* Also, *if* $R \neq 1$, *the series diverges for* $R < |x| < 1$ *and* $1 < |x| < 1/R$.

Furthermore, in §§3 and 4 we shall prove

THEOREM II: *By proper choice of the coefficients,* (1), (1′) *will define near zero any function* $f_1(x)$ *which is analytic at zero, and simultaneously define near infinity any function* $f_2(x)$ *which is analytic at infinity. If* γ *is a singularity of* $f_1(x)$ *whose absolute value is not greater than that of any other singularity of* $f_1(x)$, *and*

[1] Mathematische Werke v. 2, pp. 201–33.

[2] Leçons sur les séries de polynomes, p. 93 (1910).

[3] Leçons sur les fonctions monogènes, p. 33 (1917).

if γ' is a singularity of $f_2(x)$ whose absolute value is not less than that of any other singularity of $f_2(x)$, then the radius of convergence R of the corresponding series will be the smallest[4] *of* $|\gamma|$, $1/|\gamma'|$ *and* 1. *This expansion of $f_1(x)$ and $f_2(x)$ is moreover unique.*

The coefficients in (1) can be determined in terms of the coefficients of the Taylor's expansions of $f_1(x)$ about zero and of $f_2(x)$ about infinity by means of certain recurrence relations. These coefficients may also be determined by means of certain contour integrals, provided $f_2(x)$ vanishes at infinity, (19, §5).

In §6 the series with the particular functions (1′) is shown to be a special case of a class of expansions in series of the form (1) where $F_n(x)$ is any set of functions which are all analytic in regions including zero and infinity, and such that $F_n(x)$ has a zero of order n at the origin, and

$$F_n(1/x) = \pm x F_n(x). \tag{2}$$

The convergence characters in this general case are not as simple as for the special functions (1′), but, with one more assumption on $F_n(x)$, the coefficients can always be determined so the series will define $f_1(x)$ in some circle with zero as center, and $f_2(x)$ outside a corresponding circle.

The expansions studied in this paper bear a close relationship to those considered in a paper by Graesser,[5] who, however, expanded only a single function.

2. **Convergence character of the special series:** We investigate first the region of convergence of (1) for the special functions (1′). Let $\operatorname*{Lim\,sup}_{n\to\infty} |\alpha_n|^{1/n} = 1/R_1$. Then

$$\begin{aligned}\operatorname*{Lim\,sup}_{n\to\infty} |\alpha_n F_n(x)|^{1/n} &= |x|/R_1 \text{ if } |x| < 1,\\ &= 1/(R_1|x|) \text{ if } |x| > 1.\end{aligned}$$

Hence if $R_1 < 1$, the series of odd terms of (1) converges absolutely for $|x| < R_1$ and $|x| > 1/R_1$ and diverges for $R_1 < |x| < 1$ and for $1 < |x| < 1/R_1$; and if $R_1 \geqq 1$, it converges absolutely for all values such that $|x| \neq 1$. If R_1' is any positive number smaller than R_1 and 1, the series $\Sigma|\alpha_n F_n(R_1')|$ and $\Sigma|\alpha_n F_n(1/R_1')|$ are convergent, and moreover dominate the odd terms of (1) in the respective regions $|x| \leqq R_1'$, $|x| \geqq 1/R_1'$. Hence the odd terms of (1) converge uniformly in these regions. Corresponding statements can be made for the even terms of (1) by replacing $1/R_1$ by $1/R_2 \equiv \operatorname*{Lim\,sup}_{n\to\infty} |\beta_n|^{1/n}$ etc. From these considerations and the fact that if $R_1 < 1$,

$$\operatorname*{Lim\,sup}_{n\to\infty} |\alpha_n F_n(x)| = \infty \qquad \text{for } R_1 < |x| < 1 \text{ or } 1 < |x| < 1/R_1,$$

there follows Theorem I and also

---

[4] Or one of the smallest in case two are equal. The terms "the greatest," "the smallest," "the larger," etc. used frequently in this paper, are not intended to imply that there is only one greatest, etc.

[5] American Journal of Mathematics, v. 49, p. 577 (1927).

THEOREM III: *The radius of convergence R of* (1), (1′) *is equal to the smallest of*

$$1/\operatorname*{Lim\,sup}_{n\to\infty} |\alpha_n|^{1/n}, \qquad 1/\operatorname*{Lim\,sup}_{n\to\infty} |\beta_n|^{1/n}, \qquad \text{and } 1.$$

Suppose that $R \neq 1$, and that (1) converges absolutely at some point x_0 on either circle of convergence. Then $\Sigma\alpha_n F_n(x_0)$ converges absolutely, and

$$\operatorname*{Lim}_{n\to\infty} |\alpha_n F_n(x)/\alpha_n F_n(x_0)| \leqq 1/R \qquad \text{for } x = R, \text{ or } x = 1/R,$$

as may be shown by a consideration of the separate cases involved. Hence the series $\Sigma\,|\alpha_n F_n(R)|$ and $\Sigma\,|\alpha_n F_n(1/R)|$ converge. But these series dominate the odd terms of (1) in the respective regions $|x| \leqq R$, $|x| \geqq 1/R$. Similar statements hold for the even terms of (1), so there follows

THEOREM IV: *If* (1), (1′) *converges absolutely at a single point on one of the circles of convergence, and if $R \neq 1$, it converges absolutely and uniformly in the closed regions $|x| \leqq R$, $|x| \geqq 1/R$.*

3. **Expansion of $a_p x^p$ and $b_p x^{-p}$**: Expanding (1′) in Taylor's series about zero and infinity, for $n = 0, 1, 2, \cdots$,

$$(3)\qquad \begin{aligned} F_n(x) &= \sum_{s=0}^{\infty} c_{ns}\, x^{n+s}, && |x| < 1, \\ &= -\sum_{s=0}^{\infty} c_{ns}\, x^{-n-s-1}, && |x| > 1, \end{aligned}$$

where c_{ns} is 1 or 0 according as $s/(2n+1)$ is or is not a positive integer or zero. In particular, $c_{ns} = 0$ if $0 < s < 2n$, $n > 0$. We now write formally, for $p = 0, 1, \cdots$,

$$(4)\qquad \begin{aligned} \alpha_{p0}F_0(x) + \beta_{p0}\, x\, F_0(x) + \alpha_{p1}F_1(x) + \beta_{p1}\, x\, F_1(x) + \cdots &= a_p x^p, \ |x| < 1, \\ &= b_p x^{-p}, \ |x| > 1, \end{aligned}$$

and proceed to determine the α's and β's by expanding the terms and equating coefficients. We thus obtain the recurrence relations, for $p = 0, 1, \cdots$,

$$(5)\qquad \begin{aligned} &\alpha_{pk} = \beta_{pk} = 0 \qquad \text{for } k = 0, 1, \cdots, p-1, \\ &\alpha_{pp} = a_p, \qquad \beta_{pp} = -b_p, \end{aligned}$$

$$(6)\qquad \begin{aligned} \alpha_{p,p+k} = &-\beta_{p,p+k-1} - \alpha_{p,p+k-1}c_{p+k-1,1} - \beta_{p,p+k-2}c_{p+k-2,1} - \alpha_{p,p+k-2}c_{p+k-2,2} \\ &- \cdots - \beta_{p,p}c_{p,k-1} - \alpha_{p,p}c_{p,k}, \qquad (k = 1, 2, \cdots), \end{aligned}$$

and a corresponding relation (7) obtained by interchanging the α's and β's in (6). By means of (5), (6) and (7) the α's and β's may be determined successively, and the determination is moreover unique. In particular, for the identically zero functions, $a_p = b_p = 0$, the α's and β's all vanish.

Considering (5), (6) and (7) as now the defining equations for the α's and β's, we proceed to examine the convergence of the corresponding series (4). To

obtain a bound for the α's and β's we consider the set of positive numbers $A_{p,p+k}$, $(p = 0, 1, \cdots ; k = 1, 2, \cdots)$ defined by the relations

$$A_{pp} = 1,$$

$$A_{p,p+k} = A_{p,p+k-1} + A_{p,p+k-1}c_{p+k-1,1} + A_{p,p+k-2}c_{p+k-2,1} + A_{p,p+k-2}c_{p+k-2,2} + \cdots + A_{p,p}c_{p,k-1} + A_{p,p}c_{p,k}\,. \tag{8}$$

It is easy to show by induction that if d_p is the larger of $|a_p|$ and $|b_p|$, then for $k = 0, 1, \cdots ; p = 0, 1, \cdots$,

$$|\alpha_{p,p+k}| \leqq d_p A_{p,p+k}, \quad |\beta_{p,p+k}| \leqq d_p A_{p,p+k}\,.$$

In the left member of (8), every one of the c's of the form $c_{p+k-q,q}$; $p = 0, 1, \cdots ; k = 2, 3, \cdots ; q = 1, 2, \cdots k$, will vanish for $q < 2(p + k - q)$, $p + k - q > 0$, or certainly for $q < 2k/3$. Similarly those of form $c_{p+k-q,q-1}$; $p = 0, 1, \cdots ; k = 4, 5, \cdots ; q = 2, 3, \cdots k$, will also vanish for $q < 2k/3$. Hence for $p = 0, 1, \cdots ; k = 1, 2, \cdots$, there will be after the first term not more than $2k/3 + 2$ (or $3k$) non-vanishing terms, the largest of which cannot exceed $A_{p,p+[k/3]}$, where $[k/3]$ is the largest integer (or zero) which does not exceed $k/3$. Hence, for $p = 0, 1, \cdots ; k = 1, 2, \cdots$,

$$A_{p,p+k} \leqq A_{p,p+k-1} + 3kA_{p,p+[k/3]}\,.$$

Next consider a set of numbers $A'_{p,p+k}$ defined by the relations

$$A'_{p,p+k} = A'_{p,p+k-1} + 3kA'_{p,p+[k/3]}\,,$$

$$A'_{pp} = 1\,, \qquad p = 0, 1, \cdots ; k = 1, 2, \cdots .$$

These equations fully determine the $A'_{p,p+k}$, and moreover they are independent of p. We can therefore drop the subscript p and write

$$A_k = A'_{p,p+k}\,, \qquad k = 0, 1, \cdots ; p = 0, 1, \cdots .$$

Obviously,

$$A_{p,p+k} \leqq A_k\,, \qquad k = 0, 1, \cdots ; p = 0, 1, \cdots .$$

Consider now the function a^k, $a > 1$, $k = 0, 1, \cdots$. We have from the law of the mean,

$$a^k - a^{k-1} > a^{k-1} \log a\,,$$

so that,

$$\lim_{k\to\infty} (a^k - a^{k-1})/(3ka^{k/3}) \geqq \lim_{k\to\infty} (a^{k-1} \log a)/(3ka^{k/3}) = \infty\,.$$

Hence, for every $a > 1$ there corresponds an integer N such that

$$a^k - a^{k-1} > 3ka^{k/3} \geqq 3ka^{[k/3]}\,, \qquad k > N\,. \tag{9}$$

Now

$$A_k = A_{k-1} + 3kA_{[k/3]}, \qquad k = 1, 2, \cdots . \tag{10}$$

Let M be the largest of $A_0, A_1, \cdots, A_N$. Then

$$A_k \leqq Ma^k \text{ for } k = 0, 1, \cdots, N . \tag{11}$$

From (9), (10), and (11),

$$A_{N+1} \leqq M \left\{ a^N + 3(N+1)\, a^{[(N+1)/3]} \right\} < Ma^{N+1} .$$

Continuing in this way, we get (11) for $k = 0, 1, \cdots$; and as a consequence, $\operatorname*{Lim\,sup}_{k\to\infty} A_k^{1/k} \leqq 1$. Therefore, since $d_p A_k$ dominates $|\alpha_{p,p+k}|$ and $|\beta_{p,p+k}|$, the radius of convergence R of series (4) is unity. Due to the uniform convergence of (4), using a theorem due to Weierstrass,[6] one can now verify equation (4) by substituting the Taylor's expansions of the terms and using the relations (5), (6), and (7).

4. **Expansion of $f_1(x)$ and $f_2(x)$:** Let

$$f_1(x) = \sum_{p=0}^{\infty} a_p x^p , \qquad f_2(x) = \sum_{p=0}^{\infty} b_p x^{-p} , \tag{12}$$

and let

$$\operatorname*{Lim\,sup}_{p\to\infty} |a_p|^{1/p} = 1/R_3 , \qquad \operatorname*{Lim\,sup}_{p\to\infty} |b_p|^{1/p} = R_4 .$$

Then the singularity of $f_1(x)$ which is nearest the origin is at distance R_3 and the singularity of $f_2(x)$ which is farthest from the origin is at distance R_4.

Summing (4) for p,

$$\sum_{p=0}^{\infty} [\alpha_{pp} F_p(x) + \beta_{pp}\, x F_p(x) + \cdots] = \begin{cases} f_1(x) , & |x| < 1 \\ f_2(x) , & |x| > 1 . \end{cases} \tag{13}$$

If $|x| < 1$, (13) considered as a double series is dominated by

$$\sum_{p=0}^{\infty} M(1 - |x|)^{-1} d_p |x|^p [1 + |x| + a|x| + a|x|^2 + \cdots] , \tag{14}$$

also considered as a double series. This converges if the series

$$M(1 + |x|)(1 - |x|)^{-1} \sum_{p=0}^{\infty} d_p |x|^p \sum_{n=0}^{\infty} a^n |x|^n$$

converges. For any fixed x such that $|x| < 1$, since a is any number greater than 1, a can be chosen so that $a|x| < 1$. Hence the last series in turn converges if the series

$$M(1 + |x|)(1 - |x|)^{-1}(1 - a|x|)^{-1} \sum_{p=0}^{\infty} d_p |x|^p \tag{15}$$

---

[6] Loc. cit., p. 205.

converges. Now d_p is the larger of $|a_p|$ and $|b_p|$, so $\operatorname*{Lim\,sup}_{p\to\infty} d_p^{1/p}$ is the larger of $1/R_3$ and R_4. Hence (15) converges and (13) considered as a double series converges absolutely if $|x|$ is less than 1, R_3 and $1/R_4$. Because of this absolute convergence, the order of summation in (13) may be interchanged, giving

$$(16)\qquad f_1(x) = \alpha_0\, F_0(x) + \beta_0\, x\, F_0(x) + \cdots, \qquad |x| < 1, R_3, 1/R_4,$$

where

$$(17)\qquad \alpha_n = \sum_{p=0}^{n} \alpha_{pn}, \qquad \beta_n = \sum_{p=0}^{n} \beta_{pn}.$$

By a similar argument it may be shown that the series (16) defines $f_2(x)$ for $|x| > 1$, $1/R_3$, R_4. The uniqueness of this expansion is an immediate consequence of the fact that the α's and β's must vanish if $f_1(x) \equiv f_2(x) \equiv 0$.

5. **The determination of the coefficients by means of integrals:** The coefficients of (16) are given by (17) and the recurrence relations (5), (6) and (7) in terms of the coefficients a_p and b_p of the Taylor's expansions of $f_1(x)$ and $f_2(x)$ at zero and infinity. Let R' be any positive number less than 1, R_3, and $1/R_4$, C a circle with radius R', and C' a circle with radius $1/R'$, with centers at the origin, and t any point on C or C'. We have shown that for each number t there exist numbers $\alpha_n(t)$ and $\beta_n(t)$ such that

$$(18)\qquad \alpha_0(t)\, F_0(x) + \beta_0(t)\, x\, F_0(x) + \cdots = 1/(t - x),$$

valid for all values of x inside C and outside C'. In fact

$$\alpha_p = t^{-p-1};\, b_0 = 0,\ b_p = -t^{p-1},\ p = 1, 2, \cdots;\, d_p \leqq R'^{-p-1};$$

$$\begin{aligned}
\alpha_0(t) &= t^{-1}, & \beta_0(t) &= 0\\
\alpha_1(t) &= t^{-2} - t^{-1}, & \beta_1(t) &= 1 - t^{-1},\\
\alpha_2(t) &= t^{-3} - 1, & \beta_2(t) &= t - t^{-2},\\
\alpha_3(t) &= t^{-4} + t^{-2} - t^{-1} - t, & \beta_3(t) &= -t^{-3} - t^{-1} + 1 + t^2,\\
\alpha_4(t) &= t^{-5} + t^{-3} - t^{-2} + t^{-1} - 1 - t^2, & \beta_4(t) &= -t^{-4} - t^{-2} + t^{-1} - 1 + t + t^3,\\
&\cdots\cdots.
\end{aligned}$$

Hence for a fixed x inside C, (18) is dominated by the convergent series of positive constants obtained by substituting R'^{-p-1} for d_p in (14) and interchanging the order of summation. Therefore (18) converges uniformly with respect to t for each fixed x inside C. Similarly (18) may be shown to converge uniformly with respect to t for each fixed x outside C'. Multiplying (18) by $f_1(t)$ and $f_2(t)$, integrating over C and C' respectively, and subtracting, we get

$$\int_C \frac{f_1(t)\, dt}{t - x} - \int_{C'} \frac{f_2(t)\, dt}{t - x} = \alpha_0 F_0(x) + \beta_0\, x\, F_0(x) + \cdots,$$

where

$$\begin{aligned}\alpha_n &= \int_C f_1(t)\,\alpha_n(t)\,dt - \int_{C'} f_2(t)\,\alpha_n(t)\,dt,\\ \beta_n &= \int_C f_1(t)\,\beta_n(t)\,dt - \int_{C'} f_2(t)\,\beta_n(t)\,dt,\end{aligned} \tag{19}$$

all the integrals being taken in the counterclockwise direction; and the resulting series reduces to $f_1(x)$ if x is inside C and to $f_2(x)$ if x is outside C', provided $f_2(x)$ vanishes at infinity.

6. **The general series:** The recurrence relations (5), (6) and (7) evidently enable one to calculate the α's and β's formally, regardless of the values of the c's. These relations remain unchanged for any expansion of the form (4) even though the functions $F_n(x)$ do not have the specific form (1′), provided they have the expansions (3) about zero and infinity, converging in some regions which are independent of n, and provided that $c_{n0} = 1$. This will happen if the following conditions are satisfied:

(a) $F_n(x)$ is an analytic function whose singularities are confined to a region $N < |x| < 1/N$, where N is a properly chosen positive number *independent of* n.

(b) $F_n(x)$ has a zero of order n at the origin. We can thus without loss of generality take $c_{n0} = 1$, for all n. We assume that this has been done.

(c) $F_n(x)$ satisfies the functional relation (2). If the plus sign is used in (2) the term $-b_p$ in (5) is changed to $+b_p$, but this difference does not affect the argument.

When these three conditions are satisfied, for each n there corresponds a positive number M such that

$$|c_{ns}| < MN^{-s}. \tag{20}$$

We now make the further assumption:

(d) Such a number M exists which is *independent of* n.

The special functions (1′) evidently satisfy these four conditions. There exist other sets of functions satisfying these conditions, in fact, other sets of rational functions. For example, every set of functions of the form $P_n(x)/Q_n(x)$ where

$$P_n(x) = x^n + \lambda_1 x^{n+1} + \lambda_2 x^{n+2} + \cdots + \lambda_2 x^{n+\nu-2} + \lambda_1 x^{n+\nu-1} + x^{n+\nu},$$

$$Q_n(x) = 1 + \mu_1 x + \mu_2 x^2 + \cdots + \mu_2 x^{2n+\nu-1} + \mu_1 x^{2n+\nu} + x^{2n+\nu+1},$$

the μ's, λ's, and ν being functions of n, satisfies conditions (b) and (c). If all the μ's were zero, (a) would certainly be satisfied. If, in addition, the λ's are bounded in absolute value by a number M_1, independent of n, direct computation shows that (d) would be satisfied by choosing N as any number less than 1, and M as the larger of 1 and M_1.

We now consider *any* set of functions satisfying the above four conditions. Proceeding as in §3 we get (8) except that the c's would be replaced by their

absolute values. Substituting (20) in (8) *and assuming, as we evidently may, that* $M \geqq 1$, and letting

$$K = M(1 + N^{-1}), \qquad S = N(MN + M + 1)^{-1},$$

there is obtained

$$A_{p,\,p+k} \leqq K[A_{p,\,p+k-1} + A_{p,\,p+k-2}N^{-1} + \cdots + A_{p,\,p}N^{-k+1}]. \tag{21}$$

Starting with $A_{pp} = 1$ and applying (21) successively we get

$$A_{p,\,p+k} \leqq KS^{-k+1}.$$

From (20), $|F_{p+k}(x)| \leqq MN|x|^{p+k}(N - |x|)^{-1}, |x| < N$. Hence if $|x| < N$, the series (4) is dominated by the series

$$\sum_{k=0}^{\infty} d_p MKNS^{-k+1}|x|^{p+k}(1 + |x|)(N - |x|)^{-1} \tag{22}$$

which converges if $|x| < S$.

If S' is any positive number less than S, the series (22) is dominated for $|x| \leqq S'$ by the convergent series of constants formed by substituting S' for $|x|$. Hence (4) converges absolutely and uniformly for $|x| \leqq S'$, and represents $a_p x^p$ in this region. By a similar argument it converges absolutely and uniformly to the function $b_p x^{-p}$ for $|x| \geqq 1/S'$.

Next, summing (4) for p we get (13), which is dominated by the series (22) summed over p. Hence (13) will converge absolutely if $|x| < S$ and if the series

$$(1 + |x|)(N - |x|)^{-1}SMKN(1 - S^{-1}|x|)^{-1}\sum_{p=0}^{\infty} d_p|x|^p \tag{23}$$

converges. If R_3 and R_4 have the same meanings as in §4, and if S' is any number less than S, R_3, and $1/R_4$, then (23) is dominated for $|x| \leqq S'$ by the convergent series of constants obtained by substituting S' for $|x|$. Hence (13) converges absolutely and uniformly for $|x| \leqq S'$, and the order of summation may be interchanged, so the series (16), (17) converges absolutely and uniformly to $f_1(x)$ for $|x| \leqq S'$. Similarly, (16) converges absolutely and uniformly to $f_2(x)$ for $|x| \geqq 1/S'$.

The coefficients in (16) may be expressed in terms of contour integrals as in §5.

UNIVERSITY OF ILLINOIS, URBANA, ILLINOIS.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

THE POLYNOMIAL $F_n(x)$

By H. Bateman

(Received February 27, 1934)

1. The polynomial $F_n(x)$ is analytically useful on account of the equations

$$
(1)\qquad
\begin{aligned}
F_n\left(\frac{d}{du}\right) \operatorname{sech} u &= \operatorname{sech} u \cdot P_n(\tanh u),\\
F_n\left(\frac{d}{du}\right) u \operatorname{sech} u &= \operatorname{sech} u \cdot Q_n(\tanh u),\\
F_n\left(\frac{d}{du}\right) \operatorname{cosech} u &= \operatorname{cosech} u \cdot P_n(\coth u),\\
F_n\left(\frac{d}{du}\right) u \operatorname{cosech} u &= \operatorname{cosech} u \cdot Q_n(\coth u),
\end{aligned}
$$

wherein the notation is that usually employed in the theory of Legendre functions. The properties of the polynomial were developed originally by starting with a generating function,[1] but the fact that the foregoing equations hold simultaneously was then proved by a process of verification[2] and it seems worth while to adopt an alternative definition which leads more directly to these equations. When t is real and ν has a positive real part we may define a function F_ν (it) by means of the definite integral

$$
(2)\qquad F_\nu(it) \operatorname{sech}^2(\pi t/2) = \frac{2}{\pi^2} \int_{-\infty}^{\infty} e^{-izt} \operatorname{cosech} z \cdot Q_\nu(\coth z)\, dz .
$$

When ν is a positive integer n, this definition may be shown to be consistent with the original definition of $F_n(x)$ because we may use Fourier's inversion formula to obtain the equation

$$
(3)\qquad \operatorname{cosech} z \cdot Q_n(\coth z) = \frac{\pi}{4} \int_{-\infty}^{\infty} e^{ixz} \operatorname{sech}^2(\pi x/2)\, F_n(ix)\, dx \equiv F_n\left(\frac{d}{dz}\right) z \operatorname{cosech} z ,
$$

which agrees with the last of equations (1). This equation may be extended to certain complex values of z (in a region bounded by two parallel lines) and combined with the formula

$$
\lim_{\epsilon \to 0} i[Q_n\{\tanh(u + i\epsilon)\} - Q_n\{\tanh(u - i\epsilon)\}] = \pi P_n(\tanh u)
$$

---

[1] H. Bateman, The Tôhoku Mathematical Journal, vol. 37, p. 23 (1933).

[2] This process may be replaced by the analysis of §3 in which m is set equal to zero.

so as to give the formula for real values of u

$$(4)\quad \operatorname{sech} u \cdot P_n(\tanh u) = \frac{1}{2}\int_{-\infty}^{\infty} e^{ixu} \operatorname{sech}(\pi x/2) F_n(ix)\, dx \equiv F_n\left(\frac{d}{du}\right) \operatorname{sech} u \,;$$

this agrees with the first of the equations (1). This equation may next be inverted by Fourier's theorem so as to give the formula

$$(5)\qquad F_n(ix) \operatorname{sech}(\pi x/2) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-ixu} \operatorname{sech} u \cdot P_n(\tanh u)\,,$$

which has been used to obtain an orthogonal relation[3]

$$(6)\qquad \int_{-\infty}^{\infty} \operatorname{sech}^2(\pi x/2)\, F_m(ix)\, F_n(ix)\, dx = \begin{cases} 0, & m \neq n\,, \\ 4/\pi(2n+1), & m = n\,. \end{cases}$$

This relation and the expansion

$$(7)\qquad e^{-iax} = \sum_{n=0}^{\infty} (2n+1) \operatorname{cosech} a \cdot Q_n(\coth a)\, F_n(ix) \qquad (a \text{ and } x \text{ real})$$

quickly lead to the equation of Frobenius

$$(8)\quad \sum_{n=0}^{\infty} (2n+1) \operatorname{cosech} a \cdot \operatorname{cosech} z Q_n(\coth a) Q_n(\coth z) = (z-a) \operatorname{cosech}(z-a)\,.$$

This equation may be established with the aid of the Parseval theorem for Legendre constants.[4] Indeed, since

$$Q_n\,(\coth a) = \frac{1}{2}\int_{-1}^{1} \frac{P_n(\mu)\, d\mu}{\coth a - \mu}$$

and

$$Q_n\,(\coth z) = \frac{1}{2}\int_{-1}^{1} \frac{P_n(\mu)\, d\mu}{\coth z - \mu}\,,$$

the series represents the integral

$$\frac{1}{2} \operatorname{cosech} a \operatorname{cosech} z \int_{-1}^{1} \frac{d\mu}{(\coth a - \mu)\,(\coth z - \mu)}$$

which has the value indicated.

---

[3] H. Bateman, Proceedings of the National Academy of Sciences, vol. 20, p. 63 (1934). The relation may be modified by means of the relation $F_n(-ix) = (-)^n F_n(ix)$.

[4] The equation is derived in an entirely different way by G. Frobenius, J. für Math., vol. 73, p. 1 (1871). It is given in modern notation as an example in Whittaker and Watson's Modern Analysis, 4th Ed., Ex. 3, p. 323.

2. We now introduce the functions $P_{m,n}(z)$, $Q_{m,n}(z)$ defined by means of the equations

$$
\begin{aligned}
P_{m,n}(\tanh z) &= \cosh z \cdot F_m\left(\frac{d}{dz}\right) \operatorname{sech} z \cdot P_n(\tanh z)\,, \\
Q_{m,n}(\coth z) &= \sinh z \cdot F_m\left(\frac{d}{dz}\right) \operatorname{cosech} z \cdot Q_n(\coth z)
\end{aligned} \tag{9}
$$

These functions are symmetrical in the indices m and n because

$$
\begin{aligned}
P_{m,n}(\tanh z) &= \cosh z\, F_m\left(\frac{d}{dz}\right) F_n\left(\frac{d}{dz}\right) \operatorname{sech} z\,, \\
Q_{m,n}(\coth z) &= \sinh z\, F_m\left(\frac{d}{dz}\right) F_n\left(\frac{d}{dz}\right) \operatorname{cosech} z\,.
\end{aligned} \tag{10}
$$

Since

$$\operatorname{cosech} z \cdot Q_n(\coth z) = \frac{1}{2}\int_{-1}^{1} \frac{P_n(\mu)\cosh\xi\, d\mu}{\cosh(z-\xi)} \qquad (\mu = \tanh\xi)$$

we have

$$Q_{m,n}(\coth z) = \frac{1}{2}\sinh z \int_{-1}^{1} \frac{P_m[\tanh(z-\xi)]\, P_n(\mu)\cosh\xi\, d\mu}{\cosh(z-\xi)}\,. \tag{11}$$

When $z \to 0$ this integral gives the formula

$$\lim_{z\to 0} \operatorname{cosech} z \cdot Q_{m,n}(\coth z) = \begin{cases} 0\,, & m \neq n \\ (-)^n \dfrac{1}{2n+1}\,, & m = n \end{cases} \tag{12}$$

which is useful for finding the coefficients in an expansion of the type

$$f(z) = \sum_{n=0}^{\infty} a_n Q_n(z)\,.$$

According to the theorems of Neumann[5] and Thomé,[6] if R is the radius of convergence of the power series $\Sigma a_n z^n$ the series $\Sigma' a_n Q_n(z)$ converges outside an ellipse whose foci are the points $z = \pm 1$ and whose major axis is $R + R^{-1}$ when $R \leqq 1$ but is 2 when $R > 1$. The coefficient a_n is given by Neumann's rule

$$2\pi i\, a_n = (2n+1) \int f(z)\, P_n(z)\, dz\,, \tag{13}$$

where the integral is taken round the ellipse and the series obtained by repeated differentiations with respect to z all converge absolutely outside the ellipse and represent the corresponding derivatives of $f(z)$.

[5] C. Neumann, 'Ueber die Entwickelung einer Funktion nach Kugelfunktionen' Hallé (1862).

[6] L. W. Thomé, Journal für Mathematik, vol. 66, p. 337 (1866).

An application of the formula (12) gives the alternative rule

$$a_n = (-)^n (2n+1) \lim_{w \to 0} \left\{ F_n \left(\frac{d}{dw} \right) [f(\coth w) \operatorname{cosech} w] \right\} \tag{14}$$

which can also be derived from Neumann's rule by transforming the integral by the substitution $z = \coth w$ and making use of the formula

$$u\, F_n \left(\frac{d}{dw} \right) v - (-)^n v\, F_n \left(\frac{d}{dw} \right) u = \frac{d}{dw} X(u, v, w) .$$

The integral

$$\frac{1}{2\pi i} \int \operatorname{cosech} w \cdot dw\, F_n \left(\frac{d}{dw} \right) [f(\coth w) \operatorname{cosech} w]$$

can be evaluated by finding the residue of the integrand at the point $w = 0$.

The new rule has the advantage that it does not require the evaluation of a contour integral. It should be noticed that it gives the correct values of the coefficients in the expansions (8) and (7). It may be used also to define a set of Legendre constants of the second kind belonging to a function $f(z)$ when this function is known only for real values of z.

The new rule or a process of repeated differentiation may be used to prove the equation

$$\begin{aligned} &\operatorname{cosech} (z - a)\, Q_m [\coth (z - a)] \\ &\qquad = \operatorname{cosech} a \operatorname{cosech} z \sum_{n=0}^{\infty} (2n+1)\, Q_{m,n} (\coth z)\, Q_n (\coth a) . \end{aligned} \tag{15}$$

This may also be established by means of the Parseval theorem for Legendre series which indicates that the expansion is equal to the definite integral

$$\frac{1}{2} \int_{-1}^{1} \frac{P_m [\tanh (z - \xi)]\, d\mu}{(\cosh z - \mu \sinh z)(\cosh a - \mu \sinh a)} .$$

To evaluate this we write it in the form

$$\frac{1}{2} \int_{-\infty}^{\infty} \frac{P_m [\tanh (z - \xi)]\, d\xi}{\cosh(z - \xi) \cosh (a - \xi)}$$

and make the substitution $\xi = z - u$ which brings it to the form

$$\frac{1}{2} \int_{-\infty}^{\infty} \frac{P_m [\tanh u]\, du}{\cosh u \cosh (a - z + u)} .$$

The integral is now easily evaluated and is seen to have the value

$$\operatorname{cosech} (z - a)\, Q_m [\coth (z - a)] .$$

Writing the equation in the form

$$\sum_{n=0}^{\infty}(2n+1)\,Q_{m,n}(x)\,Q_n(y)=\frac{1}{y-x}\,Q_m\left[x+\frac{x^2-1}{y-x}\right]$$

(16)

$$=\sum_{p=0}^{\infty}\frac{1}{p\,!}\,\frac{(x^2-1)^p}{(y-x)^{p+1}}\,\frac{d^p}{dx^p}\,Q_m(x)\qquad (y\gg x>0)$$

and making use of the expansion

$$\frac{p\,!}{(y-x)^{p+1}}=\sum_{n=0}^{\infty}(2n+1)\,Q_n(y)\,\frac{d^p}{dx^p}\,P_n(x)$$

we find that

$$Q_{m,n}(x)=\sum_{p=0}^{\infty}\frac{1}{(p\,!)^2}(x^2-1)^p\,\frac{d^p}{dx^p}\,Q_m(x)\,\frac{d^p}{dx^p}\,P_n(x)$$

(17)

$$=\sum_{p=0}^{\infty}\frac{Q_m^p(x)\,P_n^p(x)}{(p\,!)^2}\,.$$

This series terminates at the n^{th} term after the first. There is also a series

$$Q_{m,n}(x)=\sum_{p=0}^{\infty}\frac{Q_n^p(x)\,P_m^p(x)}{p\,!\,p\,!}\tag{18}$$

which terminates at the m^{th} term after the first. The equivalence of the two series in the case $m=0$ leads to the interesting expansion

$$Q_n(x)=\sum_{p=0}^{n}\frac{Q_o^p(x)\,P_n^p(x)}{p\,!\,p\,!}\,.\tag{19}$$

If we put $m=n$ and let $x\to\infty$ we obtain the equation

$$(-)^n=\sum_{p=0}^{n}(-)^p\,\frac{1}{p\,!\,p\,!}\,\frac{(n+p)\,!}{(n-p)\,!}\,.\tag{20}$$

By making use of the formula[7]

$$\lim_{\epsilon\to 0}\,[i\,Q_{m,n}\{\tanh(z+i\epsilon)\}-i\,Q_{m,n}\{\tanh(z-i\epsilon)\}]=\pi\,P_{m,n}\,[\tanh z]$$

we may derive from (17) the expansion

$$P_{m,n}(x)=\sum_{p=0}^{\infty}\frac{1}{p\,!\,p\,!}\,(x^2-1)^p\,\frac{d^p}{dx^p}\,P_m(x)\,\frac{d^p}{dx^p}\,P_n(x)\,,\qquad -1\leqq x\leqq 1\tag{21}$$

[7] This is a simple consequence of formula (33) of §3.

which may also be written in the form[8]

$$P_{m,n}(x) = \sum_{p=0}^{\infty} \frac{(-)^p}{p!\,p!} P_m^p(x) P_n^p(x), \qquad -1 \leqq x \leqq 1. \tag{22}$$

This series also terminates and may be used for the calculation of $P_{m,n}(x)$. It is seen at once that $P_{m,n}(1) = 1$ and that

$$P_{1,1}(x) = 2x^2 - 1, \qquad P_{1,2}(x) = \frac{9x^3 - 7x}{2},$$

$$P_{2,2}(x) = \frac{27x^4 - 30x^2 + 5}{2}, \qquad P_{1,3}(x) = \frac{20x^4 - 21x^2 + 3}{2}, \cdots,$$

$$\int_{-1}^{1} P_{m,n}(x)\,dx = \begin{cases} 0, & m \neq n, \\ (-)^n \dfrac{2}{2n+1}, & m = n. \end{cases} \tag{23}$$

The last equation furnishes another rule for finding Legendre constants. Indeed, if

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(\tanh x), \tag{24}$$

the rule is that

$$\int_{-\infty}^{\infty} \operatorname{sech} x\,dx\, F_n\left(\frac{d}{dx}\right) \{f(x) \operatorname{sech} x\} = (-)^n \frac{2a_n}{2n+1}. \tag{25}$$

This rule, however, is not so useful as the rule given for the series of Q-functions. This latter rule is useful for the formal solution of the integral equation

$$f(z) = \int_{-1}^{1} \frac{\phi(\mu)\,d\mu}{z - \mu} \tag{26}$$

which occurs in potential theory when the potential of a thin heterogeneous rod is known at points on its axis outside the rod. If we assume that

$$\phi(\mu) = \sum_{n=0}^{\infty} a_n P_n(\mu), \tag{27}$$

the equation requires that we should have

$$f(z) = 2\sum_{n=0}^{\infty} a_n Q_n(z) \tag{28}$$

---

[8] If x lies outside the interval $(-1 \leqq x \leqq 1)$ the factor $(-)^p$ should be omitted.

and the rule gives

$$\phi(\mu) = \sum_{n=0}^{\infty} (-)^n (n + \tfrac{1}{2}) P_n(\mu) \lim_{x\to 0} \left\{F_n\left(\frac{d}{dx}\right) f(\coth x) \operatorname{cosech} x\right\}. \tag{29}$$

It is not necessary for this expansion to converge in order that the rule may be used to give a solution of the integral equation for the rule really gives the Legendre constants of the function $\phi(t)$ and some other method of summation may be used to derive the function $\phi(\mu)$ from them. Indeed, since

$$f(\coth x) \operatorname{cosech} x = \int_{-1}^{1} \frac{\phi(\mu) \cosh \alpha \, d\mu}{\cosh (x - \alpha)} \qquad (\mu = \tanh \alpha)$$

we find that when $\phi(\mu)$ is continuous

$$F_n\left(\frac{d}{dx}\right) f(\coth x) \operatorname{cosech} x = \int_{-1}^{1} \frac{\phi(\mu) \cosh \alpha \cdot d\mu}{\cosh (x - \alpha)} P_n [\tanh (x - \alpha)]$$

and so

$$\lim_{x\to 0} \left\{F_n\left(\frac{d}{dx}\right) f(\coth x) \operatorname{cosech} x\right\} = (-)^n \int_{-1}^{1} \phi(\mu) \, d\mu \, P_n(\mu) .$$

3. Let us now consider a polynomial $F_{m,n}(x)$ such that

$$F_{m,n}\left(\frac{d}{dx}\right) \operatorname{sech} x = \operatorname{sech} x \cdot P_m(\tanh x) \, P_n(\tanh x) . \tag{30}$$

We readily find that

$$F_{1,1}(x) = \tfrac{1}{2}(x^2 + 1) , \qquad F_{2,2}(x) = \tfrac{1}{32}(3x^4 + 18x^2 + 11) ,$$

$$F_{1,2}(x) = - \tfrac{1}{4}(x^3 + 3x) , \qquad F_{1,3}(x) = \tfrac{1}{48}(5x^4 + 34x^2 + 9) ,$$

and so it will be noticed that $F_{m,n}(1) = (-1)^{m+n}$.

To find the value of $F_{m,n}(d/dx)(x \operatorname{cosech} x)$ we write

$$x \operatorname{cosech} x = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\operatorname{sech} \xi \cdot d\xi}{\cosh (x - \xi)} ,$$

$$\begin{aligned} F_{m,n}\left(\frac{d}{dx}\right) x \operatorname{cosech} x &= \frac{1}{2} \int_{-\infty}^{\infty} \operatorname{sech} \xi \cdot \operatorname{sech} (x-\xi) \, P_m[\tanh (x-\xi)] \, P_n [\tanh (x-\xi)] \, d\xi \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \operatorname{sech} (x - t) \operatorname{sech} t \, P_m(\tanh t) \, P_n(\tanh t) \, dt \\ &= \frac{1}{2} \operatorname{cosech} x \int_{-1}^{1} \frac{P_m(\mu) \, P_n(\mu)}{\coth x - \mu} \frac{d\mu}{} . \end{aligned}$$

Hence, by a known formula,[9]

$$F_{m,n}\left(\frac{d}{dx}\right) x \operatorname{cosech} x = \begin{cases} \operatorname{cosech} x \cdot Q_m(\coth x) \, P_n(\coth x) , & m \geqq n , \\ \operatorname{cosech} x \cdot P_m(\coth x) \, Q_n(\coth x) , & m \leqq n . \end{cases} \tag{31}$$

[9] See, for instance, H. Bateman, Cambridge Philosophical Transactions, vol. 21, p. 179 (1909). The formula can be regarded as an old one because it is obtained immediately by a simple device used by Stieltjes in 1890. Oeuvres complètes de Thomas Jan Stieltjes, Groningen (1918), t. 2, p. 261.

We infer also that

$$(32)\qquad \begin{aligned} F_{m,n}\left(\frac{d}{dx}\right) x \operatorname{sech} x &= \operatorname{sech} x\, Q_m(\tanh x)\, P_n(\tanh x)\,, \qquad m > n \\ F_{m,n}\left(\frac{d}{dx}\right) \operatorname{cosech} x &= \operatorname{cosech} x \cdot P_m(\tanh x)\, P_n(\tanh x)\,. \end{aligned}$$

Similar analysis shows that

$$F_m\left(\frac{d}{dx}\right) F_n\left(\frac{d}{dx}\right) x \operatorname{cosech} x = \frac{1}{2} \operatorname{cosech} x \int_{-1}^{1} \frac{P_{m,n}(\theta)\, d\theta}{\coth x - \theta}.$$

Hence we have the formula

$$(33)\qquad Q_{m,n}(z) = \frac{1}{2} \int_{-1}^{1} \frac{P_{m,n}(\theta)\, d\theta}{z - \theta} \qquad (z = \coth x)\,.$$

We may also write

$$(34)\qquad \begin{aligned} F_m\left(\frac{d}{dx}\right) F_n\left(\frac{d}{dx}\right) \operatorname{cosech} x &= \operatorname{cosech} x\, P_{m,n}(\coth x)\,, \\ F_m\left(\frac{d}{dx}\right) F_n\left(\frac{d}{dx}\right) x \operatorname{sech} x &= \operatorname{sech} x\, Q_{m,n}(\tanh x)\,, \end{aligned}$$

where $Q_{m,n}$ $(\tanh x)$ is associated with $Q_{m,n}$ $(\coth x)$ in much the same way that Q_n $(\tanh x)$ is associated with Q_n $(\coth x)$.

4. Let a polynomial $E_n(x)$ be defined by means of the equation

$$(35)\qquad E_n\left(\frac{d}{du}\right) \operatorname{sech}^2 u = \operatorname{sech}^2 u\, P_n(\tanh u)\,.$$

It is readily found that

$$E_0(x) = 1\,, \qquad E_1(x) = -\frac{1}{2}x\,,$$

$$E_2(x) = \frac{1}{4}x^2\,, \qquad E_3(x) = \frac{4x - 5x^3}{48}\,.$$

Since

$$\operatorname{sech}^2 u = \frac{1}{2} \int_{-\infty}^{\infty} e^{-ixu}\, x \operatorname{cosech}(\pi x/2)\, dx\,,$$

we have

$$(36)\qquad \operatorname{sech}^2 u \cdot P_n(\tanh u) = \frac{1}{2} \int_{-\infty}^{\infty} e^{-ixu}\, x \operatorname{cosech}(\pi x/2)\, E_n(-ix)\, dx\,,$$

and Fourier's inversion formula suggests that

$$x \operatorname{cosech}(\pi x/2)\, E_n(-ix) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{ixu} \operatorname{sech}^2 u\, P_n(\tanh u)\, du\,. \tag{37}$$

Putting $\mu = \tanh u$ we have

$$G_n(-ix) = \frac{1}{2}\int_{-1}^{1}\left(\frac{1+\mu}{1-\mu}\right)^{\frac{ix}{2}} P_n(\mu)\, d\mu = \frac{1}{2}\pi x \operatorname{cosech}(\pi x/2)\, E_n(-ix)\,. \tag{38}$$

Making use of the expansion

$$P^n(\mu) = 1 - \binom{n+1}{2}\binom{2}{1}\left(\frac{1-\mu}{2}\right) + \binom{n+2}{4}\binom{4}{2}\left(\frac{1-\mu}{2}\right)^2 - \cdots$$

we obtain the series

$$E_n(2z) = 1 - \frac{1}{2}\binom{n+1}{2}\binom{2}{1}\binom{z+1}{1} + \frac{1}{3}\binom{n+2}{4}\binom{4}{2}\binom{z+2}{2} - \cdots\,. \tag{39}$$

It should be noticed that

$$G_n(0) = E_n(0) = \frac{1}{2}\int_{-1}^{1} P_n(\mu)\, d\mu = 0\,, \qquad n > 0$$

$$\begin{aligned} G_n(z+1) + G_n(z-1) &= \int_{-\infty}^{\infty} e^{-zu} \operatorname{sech} u\, P_n(\tanh u)\, du \\ &= \pi F_n(z) \sec(\pi z/2)\,, \qquad -1 < z < 1 \end{aligned} \tag{40}$$

$$e^{ixt} = \sum_{n=0}^{\infty} (2n+1)\, G_n(-ix)\, P_n(\tanh t)\,,$$

where x and t are real.

PASADENA, CAL.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

AN EXTENSION OF KRONECKER'S THEOREM

By H. M. Bacon

(Received July 28, 1933)

In a remarkable paper, "Näherungsweise ganzzahlige Auflösung linearer Gleichungen" (1884)[1] Kronecker has established some general theorems concerning non-homogeneous Diophantine inequalities. The simplest of these theorems can be stated as follows:

Given two series of real numbers, $\mu_1, \mu_2, \cdots, \mu_n$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$, and an arbitrarily small positive number ϵ, it is possible to find a set of $n + 1$ integers $g_0, g_1, \cdots, g_n$ satisfying the inequalities

$$|g_0\alpha_j - \mu_j - g_j| < \epsilon, \qquad (j = 1, 2, \cdots, n)$$

provided no relation $c_0 + c_1\alpha_1 + \cdots + c_n\alpha_n = 0$ with integral $c_0, c_1, \cdots, c_n$ can exist unless $c_0 = c_1 = \cdots = c_n = 0$. The constants $\mu_1, \mu_2, \cdots, \mu_n$ can be taken arbitrarily.

It is not difficult to show that this theorem can be put into a somewhat simpler form, as follows:

Constants $\alpha_1, \alpha_2, \cdots, \alpha_n$ being linearly independent while $\mu_1, \mu_2, \cdots, \mu_n$ are arbitrarily given real numbers, there exists a number t such that differences

$$t\alpha_j - \mu_j, \qquad (j = 1, 2, \cdots, n)$$

are in absolute value, and modulo 1, less than any preassigned positive number ϵ.

The possibility of making the above mentioned differences, modulo 1, less than any preassigned number ϵ introduces the requirement of linear independence of the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$. Recently H. Bohr and E. Landau raised the question, what condition to be imposed upon $\alpha_1, \alpha_2, \cdots, \alpha_n$ would be necessary and sufficient in order to satisfy Kronecker's inequalities, not for an *arbitrary* ϵ, but for a *given* ϵ? It is evident that the complete linear independence of these constants is too heavy a requirement when ϵ is given, and one can expect that for a given ϵ there may exist relations of the type $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$ in which the integers $c_1, c_2, \cdots, c_n$ exceed a certain limit depending upon ϵ. Although in this paper the conditions recognized as sufficient for satisfying Kronecker's inequalities do not in general coincide with the conditions found to be necessary, still the results seem to present some interest.

I. Necessary condition. A given number can be assumed to be equal to $1/N$. Suppose there exists a relation $\sum_{j=1}^{n} c_j\alpha_j = 0$ where $\sum_{j=1}^{n} |c_j| \leqq N/2$.

[1] *Leopold Kronecker's Werke*, Leipzig, 1899, III, 47–110.

Under such circumstances we can show that it is impossible to satisfy the inequalities

$$(1.1) \qquad | \alpha_j t - \mu_j - g_j | < 1/N \qquad (j = 1, 2, \cdots, n)$$

with a real t and integral g_j and appropriately chosen numbers μ_j. To this end we remark that these inequalities are equivalent to a set of equalities

$$\alpha_j t - \mu_j - g_j = \theta_j/N, \qquad | \theta_j | < 1, \qquad (j = 1, 2, \cdots, n).$$

Hence

$$\sum_{j=1}^{n} c_j\mu_j + \sum_{j=1}^{n} c_j g_j = -(1/N) \sum_{j=1}^{n} c_j\theta_j .$$

Since the numbers $\mu_1, \mu_2, \cdots, \mu_n$ are arbitrary, they can be so chosen that $\sum_{j=1}^{n} c_j\mu_j = \frac{1}{2}$, and since

$$| -(1/N) \sum_{j=1}^{n} c_j\theta_j | < (1/N)\cdot(N/2) = \tfrac{1}{2} ,$$

we come to the impossible result: $\frac{1}{2}$ + an integer is numerically less than $\frac{1}{2}$.

Therefore it is necessary for satisfying the inequalities (1.1) that

$$\sum_{j=1}^{n} | c_j | > N/2 .$$

II. **Sufficient conditions.** If there exists only one independent relation $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$, the preceding necessary condition is also sufficient. However, if more than one such relation exists, the writer has not succeeded in showing the necessary condition to be at the same time sufficient. It is possible to show that if $\sum_{j=1}^{n} | c_j | > k\cdot N$ where k is a constant, then Kronecker's inequalities will be satisfied by some real number t, and hence in general it is necessary and sufficient that $\sum_{j=1}^{n} | c_j |$ be of the order of N. In finding sufficient conditions, a method is used which is based partly upon considerations belonging to the geometry of numbers, and partly upon a generalization of Kronecker's Theorem proved by Bohr.

1. **Generalized Kronecker's Theorem.** The generalized Kronecker's Theorem may be stated as follows:

It is possible to solve the system of inequalities $| \alpha_j t - \mu_j | < \epsilon$ (mod. 1), $j = 1, 2, \cdots, n$ *for an arbitrarily given* ϵ, *if to any linear relation* $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$ *there corresponds an integral value of the sum* $c_1\mu_1 + c_2\mu_2 + \cdots + c_n\mu_n$.[2]

[2] In *Det Kongelige Danske Videnskabernes Selskat Mathematikfysiske Meddelelser*, København, 1924, VI, 8.

2. Fundamental system of linear relations. Admitting the possibility of linear relations with integral coefficients between the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$, the number of *independent* relations of this type cannot be greater than $n - 1$. Suppose there are exactly s ($\leqq n - 1$) such relations. The system of s independent relations can be selected in various ways. We shall suppose that of all these systems we select one

$$\sum_{j=1}^{n} a_{ij}\,\alpha_j = 0 \qquad (i = 1, 2, \cdots, s)$$

with the greatest common divisor[3] of the determinants of order s in the matrix

$$\left\| \begin{matrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ a_{s1} & \cdots & a_{sn} \end{matrix} \right\| \tag{2.1}$$

a minimum. For convenience we can designate such a system of s independent relations as a *fundamental system.* The important property of a fundamental system is that in any other linear relation $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$ the integers $c_1, c_2, \cdots, c_n$ can be expressed as linear combinations with integral coefficients of the elements of the $1^{\text{st}}, 2^{\text{nd}}, \cdots, n^{\text{th}}$ columns of the matrix (2.1). In other words, $c_j = \sum_{i=1}^{s} h_i a_{ij}$, $(j = 1, 2, \cdots, n)$, with integral h_i.

This remark shows that in the generalized Kronecker's Theorem we need not consider all possible linear relations between the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$. It suffices that the conditions for the validity of this theorem be fulfilled for only s independent relations forming a fundamental system.

3. Sufficient condition in case of one relation between the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$. In case there exists a single independent relation

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$$

it is sufficient for the existence of a real number t satisfying the inequalities

$$| \alpha_j t - \mu_j | < \epsilon = 1/N \text{ (mod. 1)} \qquad (j = 1, 2, \cdots, n)$$

to have $\sum_{i=1}^{n} | c_i | > N/2$. For consider, in a space of n dimensions, the planes

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = h \tag{3.1}$$

where h takes all integral values $0, \pm 1, \pm 2, \cdots$. If $(\mu_1, \mu_2, \cdots, \mu_n)$ is an arbitrary point, and if $(\nu_1, \nu_2, \cdots, \nu_n)$ is a point on one of these planes, then

$$\sum_{j=1}^{n} c_j\mu_j = a\,, \qquad \sum_{j=1}^{n} c_j\nu_j = h\,, \qquad \sum_{j=1}^{n} c_j(\mu_j - \nu_j) = a - h$$

[3] It can be shown that this greatest common divisor is 1, but this is irrelevant to our purpose.

where a is some constant. We consider a generalized distance given by $\mathrm{Max}\,\{|\mu_1 - \nu_1|, \cdots, |\mu_n - \nu_n|\}$. If we take a point $(\nu_1, \nu_2, \cdots, \nu_n)$ where ν_j is determined by the equations

$$\mu_j - \nu_j = \frac{(a - h)\,\mathrm{sgn}\,c_j}{\sum_{i=1}^{n} |c_i|}, \qquad (j = 1, 2, \cdots, n)$$

it will be seen to lie upon a plane (3.1). By the generalized Kronecker's Theorem, for any arbitrary σ, $|\alpha_j t - \nu_j| < \sigma$ (mod. 1). By an appropriate choice of h, we may make

$$|\mu_j - \nu_j| \leqq 1/2 \sum_{i=1}^{n} |c_i|, \qquad (j = 1, 2, \cdots, n).$$

Since $\sum_{i=1}^{n} |c_i| > N/2$ by hypothesis, and since σ may be chosen arbitrarily small, we have

$$|\alpha_j t - \mu_j| < \sigma + 1/2 \sum_{i=1}^{n} |c_i| < 1/N \quad (\mathrm{mod.}\ 1).$$

Hence we may state the theorem:

In case of the existence of only one independent relation

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0,$$

a necessary and sufficient condition for the existence of a real number t satisfying the inequalities $|\alpha_j t - \mu_j| < \epsilon = 1/N$ (mod. 1) is that $c_1, c_2, \cdots, c_n$ satisfy the inequality $\sum_{i=1}^{n} |c_i| > N/2$.

If we try to make use of a generalized distance as defined above for the general case of s ($<n$) independent relations $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$, it becomes difficult to achieve any satisfactory results. For this reason it seems advisable to replace this generalized distance by the ordinary Euclidean distance and so make possible the use of known facts concerning quadratic forms.

4. **A geometrical problem.** The preceding considerations lead to the following geometrical problem: Given, in a space of n dimensions, the system of manifolds of $n - s$ dimensions determined by

$$(S) \qquad \sum_{j=1}^{n} a_{ij} x_j = h_i \qquad (i = 1, 2, \cdots, s)$$

where the h_i run over all integral values. Let δ represent the distance from a point $(\mu_1, \mu_2, \cdots, \mu_n)$ to a manifold of this system determined by fixed values of h_i. If, now, the h_i are given all possible integral values, what is the upper

bound of the minimum value of δ? Since the manifolds are supposed to be of $n - s$ dimensions, not all determinants of order s in the matrix

$$(T) \qquad \begin{Vmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ a_{s1} & \cdots & a_{sn} \end{Vmatrix}$$

vanish.

We first find the analytical expression for the distance δ of a point $(\mu_1, \mu_2, \cdots, \mu_n)$ from a determined manifold corresponding to given integral values of the h_i. Let us introduce the quadratic form

$$f(\lambda_1, \lambda_2, \cdots, \lambda_s) = \sum_{i=1}^{s} \sum_{j=1}^{s} k_{ij}\lambda_i\lambda_j$$

where $k_{ij} = \sum_{\alpha=1}^{n} a_{i\alpha}a_{j\alpha}$, $(i, j = 1, 2, \cdots, s)$. The determinant D of the form f is the sum of the squares of all the determinants of order s in the matrix (T), and hence is different from zero. If, now, we represent by d_i the differences

$$d_i = h_i - \nu_i$$

where

$$\nu_i = \sum_{j=1}^{n} a_{ij}\mu_j$$

we have for δ^2 the following expression:

$$(4.1) \qquad \delta^2 = \frac{F(d_1, d_2, \cdots, d_s)}{D}$$

where F is the adjoint form to f. Since $d_1, d_2, \cdots, d_s$ for a given manifold and a given point are known, the formula (4.1) gives the square of the shortest distance δ from this point to the manifold. To find the shortest distance from this point to all the manifolds of the system (S), we must consider the integers $h_1, h_2, \cdots, h_s$ as variable, and under such conditions seek the minimum of the right hand member of equation (4.1).

5. **Upper limit for the distance** δ. Let (e_{ij}) be a unimodular substitution with integral coefficients. This substitution transforms the quadratic form F into the quadratic form

$$F' = \sum_{i=1}^{s} \sum_{j=1}^{s} A_{ij}\, d'_i d'_j$$

in the new variables $d'_1, d'_2, \cdots, d'_s$. We can suppose the substitution (e_{ij}) chosen so as to produce a *reduced* form F'. According to the theory of reduction

of positive quadratic forms developed by Minkowski,[4] in a *reduced* form the coefficients satisfy, necessarily, the following inequalities:

$$0 < A_{11} \leqq A_{22} \leqq \cdots \leqq A_{ss}$$

and

$$2 \cdot | A_{ij} | \leqq A_{ii} \leqq A_{ss} \qquad (i < j) .$$

Furthermore, if Δ is the determinant of the form, there exists a constant Γ_s such that

$$A_{11} \cdot A_{22} \cdots A_{ss} \leqq \Gamma_s \cdot \Delta .$$

Bieberbach and Schur[5] established that one can take

$$\Gamma_s = \left(\frac{125}{48}\right)^{\frac{s^3 - s}{6}} .$$

We now have

$$\delta^2 = \frac{F(d_1, d_2, \cdots, d_s)}{D} = \frac{\sum_{i=1}^{s} \sum_{j=1}^{s} A_{ij}(h'_i - \nu'_i)(h'_j - \nu'_j)}{D} .$$

Integers $h'_1, h'_2, \cdots, h'_s$ are arbitrary and can be determined so as to have

$$| h'_i - \nu'_i | \leqq \tfrac{1}{2} , \qquad (i = 1, 2, \cdots, s) .$$

Then the distance δ of the point $(\mu_1, \mu_2, \cdots, \mu_n)$ from the manifolds (S) satisfies the inequality

$$\delta^2 \leqq \frac{\sum_{i=1}^{s} \sum_{j=1}^{s} | A_{ij} |}{4\,D} \leqq \frac{(s^2 + s) \,.\, A_{ss}}{8\,D}$$

where the right hand member does not depend upon the point $(\mu_1, \mu_2, \cdots, \mu_n)$.

6. **Generalized Hadamard's inequality.** Denoting by A_i the cofactor of X_i in the matrix

$$\begin{Vmatrix} X_1 & \cdots & X_s \\ x_1^{(1)} & \cdots & x_s^{(1)} \\ \cdot & \cdots & \cdot \\ x_1^{(s-1)} & \cdots & x_s^{(s-1)} \end{Vmatrix}$$

[4] Hermann Minkowski, "Diskontinuitätsbereich für arithmetische Äquivalenz" in his *Gesammelte Abhandlungen*, Leipzig und Berlin, 1911, II, 53–100.

[5] Bieberbach und Schur, "Über die Minkowskische Reduktionstheorie der positiven quadratische Formen" in *Sitzungsberichte der Preussischen Akademie des Wissenschaften* Berlin, 13 Dezember, 1928, 510–535.

we have

$$(6.1)\quad F(A_1, \cdots, A_s) \leqq f(x_1^{(1)}, \cdots, x_s^{(1)}) \cdots f(x_1^{(s-1)}, \cdots, x_s^{(s-1)}).$$

Here f is a positive quadratic form in s variables, and F is its adjoint. Since the adjoint of F is equal to $D^{s-2}\cdot f$, we can derive from (6.1) the similar inequality

$$(6.2)\quad D^{s-2}\cdot f(A_1, \cdots, A_s) \leqq F(x_1^{(1)}, \cdots, x_s^{(1)}) \cdots F(x_1^{(s-1)}, \cdots, x_s^{(s-1)}).$$

The proof of these inequalities will be given in an appendix to this paper.

7. **Solution of the geometrical problem.** Suppose that in the matrix

$$\left\| \begin{matrix} x_1^{(1)} & \cdots & x_s^{(1)} \\ \cdot & \cdots & \cdot \\ x_1^{(s-1)} & \cdots & x_s^{(s-1)} \end{matrix} \right\|$$

the 1^{st}, 2^{nd}, $\cdots$, $(s-1)^{th}$ rows are identical correspondingly with the 1^{st}, 2^{nd}, $\cdots$, $(s-1)^{th}$ columns of the matrix of the unimodular substitution (e_{ij}). Then $F(x_1^{(i)}, \cdots, x_s^{(i)}) = A_{ii}$ $(i = 1, 2, \cdots, (s-1))$, and hence inequality (6.2) of paragraph 6 gives

$$D^{s-2}\cdot f(A_1, \cdots, A_s) \leqq A_{11}\cdot A_{22} \cdots A_{s-1,s-1}.$$

Since $A_1, A_2, \cdots, A_s$ do not vanish simultaneously, $f(A_1, \cdots, A_s)$ is not less than the minimum M of the form $f(\lambda_1, \cdots, \lambda_s)$ for integral values of $\lambda_1, \lambda_2, \cdots, \lambda_s$ not vanishing simultaneously. It follows from this remark that

$$D^{s-2}\cdot M \leqq A_{11}\cdot A_{22} \cdots A_{s-1,s-1}.$$

On the other hand

$$A_{11}\cdot A_{22} \cdots A_{ss} < \Gamma_s\cdot D^{s-1},$$

hence

$$D^{s-2}\cdot M\cdot A_{ss} < \Gamma_s\cdot D^{s-1},$$

$$A_{ss} < \Gamma_s\cdot D/M,$$

and

$$\delta^2 < \Gamma_s\cdot\frac{s(s+1)}{8M}.$$

8. **Final conclusions.** Let us suppose that the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$ are connected by s $(< n)$ independent relations. As we have shown, there are always s independent and *fundamental* relations

$$\sum_{j=1}^{n} a_{ij}\alpha_j = 0, \qquad (i = 1, 2, \cdots, s)$$

so that in any relation $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$ the integral coefficients $c_1, c_2, \cdots, c_n$ can be expressed as

$$c_j = \sum_{i=1}^{s} \lambda_i a_{ij}, \qquad (j = 1, 2, \cdots, n)$$

where $\lambda_1, \lambda_2, \cdots, \lambda_s$ are integers. Hence

$$c_1^2 + c_2^2 + \cdots + c_n^2 = f(\lambda_1, \lambda_2, \cdots, \lambda_s),$$

and the minimum of the quadratic form f is simply the sum of squares of the coefficients in some relation between the constants $\alpha_1, \alpha_2, \cdots, \alpha_n$. The geometrical problem solved in the preceding paragraph shows the existence of a point $(\nu_1, \nu_2, \cdots, \nu_n)$ belonging to one of the manifolds (S) whose distance from the point $(\mu_1, \mu_2, \cdots, \mu_n)$ satisfies the inequality

$$\delta < \sqrt{\frac{\Gamma_s \cdot s(s+1)}{8M}} \leqq \sqrt{\frac{\Gamma_n \cdot n(n-1)}{8M}} = \frac{K_n}{\sqrt{M}}.$$

For the point $(\nu_1, \nu_2, \cdots, \nu_n)$, by the generalized Kronecker's Theorem, the inequalities

$$|\alpha_j t - \nu_j| < \epsilon \pmod{1}, \qquad (j = 1, 2, \cdots, n)$$

have a solution in t for an arbitrarily given ϵ. On the other hand

$$|\mu_j - \nu_j| \leqq \delta < \frac{K_n}{\sqrt{M}},$$

hence, for the same value of t,

$$|\alpha_j t - \mu_j| < \epsilon + \frac{K_n}{\sqrt{M}} \pmod{1},$$

or, taking $\epsilon = K_n/\sqrt{M}$,

$$|\alpha_j t - \mu_j| < 2K_n/\sqrt{M} \pmod{1}, \qquad (j = 1, 2, \cdots, n).$$

If $2K_n/\sqrt{M} < 1/N$, which is equivalent to

$$\sum_{j=1}^{n} c_j^2 = M > (2K_n)^2 \cdot N^2,$$

where N is a given number, we conclude that there exists a real number t satisfying the system of inequalities

$$|\alpha_j t - \mu_j| < 1/N \pmod{1} \qquad (j = 1, 2, \cdots, n).$$

Now, if $\sum_{j=1}^{n} |c_j| > 2\sqrt{n}\, K_n \cdot N$, then $\sum_{j=1}^{n} c_j^2 > (2K_n)^2 \cdot N^2$, and we may state the following theorem:

A sufficient condition for the existence of a real number t satisfying the system of inequalities

$$|\alpha_j t - \mu_j| < \epsilon = 1/N \text{ (mod. 1)} \qquad j = 1, 2, \cdots, n$$

is that in all of the two or more independent relations

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0,$$

the numbers $c_1, c_2, \cdots, c_n$ *satisfy the inequality*

$$\sum_{j=1}^{n} |c_j| > 2\sqrt{n}\cdot K_n \cdot N = \left(\frac{125}{48}\right)^{\frac{n^3-n}{12}} \cdot n \cdot \sqrt{\frac{n-1}{4}} \cdot N$$
$$> \left(\frac{125}{48}\right)^{\frac{n^3-n}{12}} \cdot (n-1)^{\frac{3}{2}} \cdot \frac{N}{2}.$$

III. **Appendix.** It remains to prove the inequality (6.1) of paragraph 6. Let

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

be a definite quadratic form in n variables, and (A) a matrix

$$\text{(A)} \qquad \left\| \begin{matrix} x_{11} & \cdots & x_{1n} \\ \cdot & \cdots & \cdot \\ x_{n-1,1} & \cdots & x_{n-1,n} \end{matrix} \right\|$$

containing $n - 1$ rows each consisting of n arbitrary real numbers. Denoting by $X_{hk} = \sum_{j=1}^{n} a_{kj}\, x_{hj}$, we construct another matrix (B)

$$\text{(B)} \qquad \left\| \begin{matrix} X_{11} & \cdots & X_{1n} \\ \cdot & \cdots & \cdot \\ X_{n-1,1} & \cdots & X_{n-1,n} \end{matrix} \right\|.$$

We compose matrices (A) and (B) by rows, thus obtaining a matrix out of which we can form a determinant (C)

$$\text{(C)} \qquad \left| \begin{matrix} f(x_{11}, \cdots, x_{1n}) & (1,2) & \cdots & (1, n-1) \\ (2,1) & f(x_{21}, \cdots, x_{2n}) & \cdots & (2, n-1) \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ (n-1, 1) & (n-1, 2) & \cdots & f(x_{n-1,1}, \cdots, x_{n-1,n}) \end{matrix} \right|$$

where the symbols (h, k) stand for the sums $\sum_{j=1}^{n} X_{hj}\, x_{kj}$. These sums are evidently polars of the quadratic form f. On the other hand, the same determinant (C) can be shown to be equal to $F(B_1, B_2, \cdots, B_n)$ where F is a form adjoint to f and $B_1, B_2, \cdots, B_n$ are the cofactors of $\xi_1, \xi_2, \cdots, \xi_n$ in the determinant[6]

$$\left| \begin{matrix} \xi_1 & \cdots & \xi_n \\ x_{11} & \cdots & x_{1n} \\ \cdot & \cdots & \cdot \\ x_{n-1,1} & \cdots & x_{n-1,n} \end{matrix} \right|.$$

[6] See Charles Hermite, "Lettres de M. Hermite à M. Jacobi sur différents objets de la Théorie des Nombres" (Première Lettre) in *Oeuvres de Charles Hermite*, Paris, 1905, I, 102.

By a linear transformation the form f can be transformed into the sum of squares

$$f(x_1, x_2, \cdots x_n) = \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2$$

of the new variables $\xi_1, \xi_2, \cdots, \xi_n$. Let $\xi_{h1}, \xi_{h2}, \cdots, \xi_{hn}$ be the values of these variables corresponding to $x_{h1}, x_{h2}, \cdots, x_{hn}$. Since the polars are covariants we shall also have

$$(h, k) = \xi_{h1}\cdot\xi_{k1} + \xi_{h2}\cdot\xi_{k2} + \cdots + \xi_{hn}\cdot\xi_{kn}.$$

Thus the determinant (C) expressed in the new variables will be

$$\text{(D)} \qquad \begin{vmatrix} \sum_{j=1}^{n} \xi_{1j}^2 & \sum_{j=1}^{n} \xi_{1j}\cdot\xi_{2j} & \cdots & \sum_{j=1}^{n} \xi_{1j}\cdot\xi_{n-1,j} \\ \sum_{j=1}^{n} \xi_{2j}\xi_{1j} & \sum_{j=1}^{n} \xi_{2j}^2 & \cdots & \sum_{j=1}^{n} \xi_{2j}\cdot\xi_{n-1,j} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ \sum_{j=1}^{n} \xi_{n-1,j}\cdot\xi_{1j} & \sum_{j=1}^{n} \xi_{n-1,j}\cdot\xi_{2j} & \cdots & \sum_{j=1}^{n} \xi_{n-1,j}^2 \end{vmatrix}.$$

Now we can make an orthogonal transformation of the variables $\xi_1, \xi_2, \cdots, \xi_n$ into new variables $\eta_1, \eta_2, \cdots, \eta_n$ so that $\eta_{1n}, \eta_{2n}, \cdots, \eta_{n-1,n}$ all vanish. The preceding determinant (D) expressed in the new variables will be

$$\begin{vmatrix} \sum_{j=1}^{n-1} \eta_{1j}^2 & \sum_{j=1}^{n-1} \eta_{1j}\cdot\eta_{2j} & \cdots & \sum_{j=1}^{n-1} \eta_{1j}\cdot\eta_{n-1,j} \\ \sum_{j=1}^{n-1} \eta_{2j}\cdot\eta_{1j} & \sum_{j=1}^{n-1} \eta_{2j}^2 & \cdots & \sum_{j=1}^{n-1} \eta_{2j}\cdot\eta_{n-1,j} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ \sum_{j=1}^{n-1} \eta_{n-1,j}\eta_{1j} & \sum_{j=1}^{n-1} \eta_{n-1,j}\cdot\eta_{2j} & \cdots & \sum_{j=1}^{n-1} \eta_{n-1,j}^2 \end{vmatrix}$$

and this determinant is simply the square

$$\begin{vmatrix} \eta_{11} & \cdots & \eta_{1,n-1} \\ \cdot & \cdots & \cdot \\ \eta_{n-1,1} & \cdots & \eta_{n-1,n-1} \end{vmatrix}^2.$$

Thus we have $F(B_1, B_2, \cdots, B_n) =$

$$\begin{vmatrix} \eta_{11} & \cdots & \eta_{1,n-1} \\ \cdot & \cdots & \cdot \\ \eta_{n-1,1} & \cdots & \eta_{n-1,n-1} \end{vmatrix}^2.$$

Now, by a well known theorem due to Hadamard,

$$\begin{vmatrix} \eta_{11} & \cdots & \eta_{1,\,n-1} \\ \cdot & \cdots & \cdot \\ \eta_{n-1,\,1} & \cdots & \eta_{n-1,\,n-1} \end{vmatrix}^2$$

$$\leqq (\eta_{11}^2 + \cdots + \eta_{1,\,n-1}^2)\cdot(\eta_{21}^2 + \cdots + \eta_{2,\,n-1}^2) \cdots (\eta_{n-1,1}^2 + \cdots + \eta_{n-1,\,n-1}^2).$$

Hence

$$F(B_1, \cdots, B_n) \leqq f(x_{11}, \cdots, x_{1n}) \cdots f(x_{n-1,\,1}, \cdots, x_{n-1,\,n})$$

since

$$\eta_{h1}^2 + \eta_{h2}^2 + \cdots + \eta_{h,\,n-1}^2 = \xi_{h1}^2 + \xi_{h2}^2 + \cdots + \xi_{hn}^2 = f(x_{h1}, \cdots, x_{hn})$$

$$(h = 1, 2, \cdots, (n-1)).$$

Hence the inequality (6.1) in paragraph 6 is proved.

STANFORD UNIVERSITY.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

ON THE METRICAL TRANSITIVITY OF THE GEODESICS ON CLOSED SURFACES OF CONSTANT NEGATIVE CURVATURE

By Gustav A. Hedlund[1]

(Received March 8, 1934)

1. **Introduction.** A dynamical system is *regionally transitive* if there exists a motion such that the points of the phase-space defined by this motion constitute a set which is everywhere dense in the phase-space. Such a motion is called *transitive*. The existence of transitive geodesics on closed surfaces of genus greater than one and of negative curvature has been known for some time.[2] More recently the existence of such geodesics under considerably less stringent hypotheses has been shown by Morse.[3]

If measure is defined in the phase-space, questions concerning the measure of the various types of motion arise. It is known[4] that in the case of closed orientable surfaces of constant negative curvature the totality of transitive (non-special) geodesics determine a point set of the phase-space which is of the measure of the phase-space. This result has not been extended to the case of non-constant negative curvature.

The system is said to be *metrically transitive* if any set of motions which defines a measurable set of points of the phase-space is either of measure zero or of the measure of the phase-space. An example of a dynamical system with one degree of freedom which is metrically transitive has been given by Birkhoff and Smith.[5]

In this paper it will be shown that the systems of geodesics on a certain set

[1] National Research Fellow.

[2] A case of constant negative curvature is considered by J. Nielsen, Om geodaetiske Linier i lukkede Mangfoldigheder med konstant negativ Krumming, Matematisk Tidskrift, 1925, p. 37. A proof for a more general Fuchsian group is given by Koebe, Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen, IV, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1929, p. 431, and the extension to the case of nonconstant curvature is simple. A proof based on a symbolic method due to Morse is indicated by Birkhoff, Dynamical Systems, American Mathematical Society Colloquium Publications, vol. IX, pp. 238–248.

[3] Morse, Does instability imply transitivity?, Proceedings of the National Academy of Sciences, January, 1934.

[4] Not specifically stated, but obviously following from the paper of P. J. Myrberg, Ein Approximationssatz für die Fuchsschen Gruppen, Acta Mathematica, vol. 57 (1931), pp. 389–409. See also G. A. Hedlund, On the measure of the non-special geodesics on a surface of constant negative curvature, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 345–348.

[5] G. D. Birkhoff and P. A. Smith, Structure analysis of surface transformations, Journal de Mathématiques, vol. 7.

of surfaces which are closed, orientable, of arbitrary genus greater than one, and of constant negative curvature are metrically transitive.

2. **Two-dimensional Riemannian manifolds.** Let M be a two-dimensional manifold as defined by Weyl.[6] Terms referring to point sets on M will be used as defined by Weyl. If M is triangulated, it is a *two-dimensional topological manifold* (a surface, according to Weyl).

If D is a domain of M such that its points are points of two neighborhoods N_P and N_Q, to D there will correspond two plane domains by the correspondence of N_P and N_Q with plane circles. A map of one of these plane domains, D_1, on the other, D_2, will be determined by defining corresponding points as those which correspond to the same point of D. If among the neighborhoods of points of M, a set of neighborhoods, N'_P, at least one for each point of M, can be so chosen that the preceding map of D_1 on D_2 is directly conformal, the two-dimensional topological manifold is a *Riemann surface.*[7] The terms *class* and *analytic* as applied to a curve on M are defined in terms of the plane curve segments which are obtained by the correspondences of the preceding paragraph. The consistency of the definitions is easily proved. The *angle at which two directed curves of Class C' intersect* at a point on M is defined to within multiples of 2π and sign as the angle at which the corresponding plane directed curve segments intersect.

The differential element of arc-length is introduced into M in the following manner. Let P be a point of M, N'_P a neighborhood of P and (x, y) the Cartesian coordinates of the plane containing the circle C, the interior of which is in one-to-one correspondence with N'_P. It will be assumed that to P there corresponds a function $\lambda_P(x, y)$ which is of Class C^3 and positive in the interior of C and the differential element of arc-length for the neighborhood N'_P will be defined as

$$ds^2 = \lambda_P(x, y)\,[dx^2 + dy^2]\,. \tag{2.1}$$

If the distance element thereby defined is to be consistent, the functions $\lambda_P(x, y)$ are subject to further conditions. If D is a domain of M such that its points lie in both of the neighborhoods N'_P and N'_Q, there will be two plane domains D_1 and D_2 corresponding to D and a conformal map of one into the other will be defined. Let the Cartesian coordinates in the plane containing D_1 be (x, y) and (u, v) those in the plane containing D_2. Let the conformal transformation taking D_2 into D_1 be given by

$$x = x(u, v)\,, \qquad y = y(u, v)\,. \tag{2.2}$$

---

[6] Weyl, Die Idee der Riemannsche Fläche, (1923), p. 16.

[7] A Riemannian manifold according to Koebe, Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen I, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1927, pp. 164–196. The Riemannian manifold of Koebe is more general than the Riemann surface, for inverse conformal maps are admitted.

The transformation (2.2) applied to the differential form (2.1) yields a differential form

$$ds^2 = \lambda'_P(u, v)\,[du^2 + dv^2], \tag{2.3}$$

where $\lambda'_P(u, v)$ is defined at least in the domain D_2. The further condition is now imposed that this function $\lambda'_P(u, v)$ must in D_2 be identical with the function $\lambda_Q(u, v)$ associated with the neighborhood N'_Q. A Riemann surface with ds^2 defined as here will be called a *two-dimensional Riemannian manifold* M_R.

The geodesics on M_R are the curves on M_R corresponding to the plane geodesics determined by (2.1).

3. **The Riemannian manifold under consideration.** Let U denote the unit circle in the complex plane and let Ψ be its interior. The geodesics in Ψ corresponding to the differential form

$$ds^2 = 4(1 - z\bar{z})^{-2}\,|\,dz\,|^2 = 4(1 - x^2 - y^2)^{-2}\,(dx^2 + dy^2) \tag{3.1}$$

are arcs of circles orthogonal to U, and if angle is defined in the ordinary sense a well-known non-euclidean geometry is defined. The Gaussian curvature is -1. The differential form (3.1) is invariant under the linear fractional transformations taking U and its interior into U and its interior, respectively. The geodesics in Ψ corresponding to (3.1) will be called *NE (non-euclidean) straight lines.* Either of the parts into which a NE straight line is divided by a point will be called a *NE ray.* Given two points P_1 and P_2 in Ψ, there is a unique NE line segment, σ, joining P_1 and P_2 and the *NE distance* between P_1 and P_2 will be defined as $\int_\sigma ds$ where ds is given by (3.1).

Let F be the Fuchsian group defined by Nielsen.[8] The fundamental region of F can be taken as a region S_0 bounded by $4p(p > 1)$ arcs of geodesics, where these arcs are all equal, symmetrically situated about the origin, and congruent in pairs. The transformations of F are hyperbolic transformations with fixed points on U, and the transformations carrying the sides of S_0 into one another form a set of generators for F. The interior angle at each vertex is $\pi/2p$ and the vertices constitute a single cycle.

Let q be a point of Ψ and γ the set of points of Ψ which are congruent to q under the transformations of the group F. This set γ can be considered as a "point" and denoted by Q. To define a neighborhood of such a point Q, let n_q be the set of points in Ψ at NE distance from q less than a positive constant which is not as great as the minimum NE distance between congruent points in Ψ. Each of the points q' in n_q determines a point Q' and the totality of such points Q' is defined as a neighborhood of Q. A neighborhood of Q is evidently in one-to-one correspondence with the interior of a plane circle. It is easily verified that the totality of such points Q obtained by letting q be any point

[8] J. Nielsen. Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen, Acta Mathematica, vol. 50 (1927), pp. 189–358. Only a knowledge of pp. 189–231 will be assumed here. This paper will hereafter be referred to as NN.

of Ψ constitutes a two-dimensional manifold as defined in the preceding paragraph. It can be triangulated and it is a closed orientable manifold of genus p. It moreover satisfies the conditions imposed on a Riemann surface, and if the differential form (3.1) is introduced, a Riemannian manifold m_R is defined. This is the Riemannian manifold considered in this paper.

A point of M_R is *represented* by a infinite set of mutually congruent points in Ψ. A point of M_R and a direction at this point is *represented* by an infinite set of mutually congruent elements in Ψ. A geodesic in M_R is *represented* by an infinite set of mutually congruent geodesics in Ψ.

4. **The Nielsen development.** By a certain process due to Nielsen (NN pp. 211–217), the unit circle is divided into $4p$ intervals, these in turn into subintervals, and the process is repeated indefinitely. The $4p$ intervals into which U is divided will be called the intervals of the first stage. With each interval of the first stage is associated an element of the set $a_i, b_i, a_i^{-1}, b_i^{-1}, i = 1, 2, \cdots p$, as well as an oriented geodesic segment of the net n, the segment emanating from the origin and bearing the same symbol as the corresponding interval. The intervals into which the intervals of the first stage are divided are the intervals of the second stage, and with each is associated a sequence of two elements. With each interval of the second stage is also associated a directed path made up of two segments of the net n, with initial point at the origin and such that the pair of symbols of these segments taken in order beginning with that incident with the origin is the sequence of elements associated with the given interval. In general, an interval δ_m of the m^{th} stage is designated by a sequence of m elements $c_1c_2 \cdots c_m$, the *symbol* of the interval δ_m. The oriented broken path $c_1c_2 \cdots c_m$ of the net n, with the origin as initial point, is also associated with δ_m, and it bears an important geometrical relation to δ_m. Let P be the terminal point of the path $c_1c_2 \cdots c_{m-1}$. Then the interval δ_m subtends through NE rays an angle at P which in general is $\pi/2p$, and is never greater than this. If δ_m is not an end subinterval of the interval of stage $(m - 1)$ in which it is contained, the angle subtended is exactly $\pi/2p$ and the NE rays form equal angles with the segment c_m of the path $c_1 \cdots c_m$.

As m increases, the lengths of the intervals of the m^{th} stage shrink uniformly to zero, and hence if intervals are considered closed, with each point of U is associated an infinite sequence of intervals (in some cases two), each contained in the preceding, and each containing the given point of U. Thus at least one infinite symbol is associated with a point of U. Except for a denumerable set on U, the infinite symbol is uniquely determined, and for each point of the exceptional set there are two symbols. With each point of U there is also associated an oriented broken path made up of an infinite sequence of segments of the net n, the path having its initial point at the origin, and being unique, except for the exceptional set, when there are two.

The sequences which occur in the finite and infinite symbols associated with

the intervals and points of U are not arbitrary. The generators of the group F satisfy the relation

$$R \equiv \prod_{i=1}^{p} a_i b_i a_i^{-1} b_i^{-1} = 1 .$$

The cyclic orders represented by R and R^{-1} will be designated by O_1 and O_2, respectively.[9] We note that the sequence R is the sequence of the sides of a polygon of the net n taken in the clockwise sense. The cyclic order represented by

$$\bar{R} \equiv \prod_{i=1}^{p} a_i b_i^{-1} a_i^{-1} b_i$$

will be designated by O_4. The order O_4 is the cyclic order of the segments of the net n emanating from a vertex and taken in the counterclockwise sense. If the cyclic order $\bar{R}$ is reversed, the resulting cyclic order will be designated by O_3.

A sequence of $2p$ elements in the order O_1 or O_2 will be called a *direct* or an *inverse semicycle*, respectively. The finite symbols of the intervals of the various stages and the infinite symbols associated with the points of U are subject to the following conditions (NN pp. 216–217):

A. *No element (generator) immediately follows its inverse.*

B. *There are no more than $2p$ consecutive elements in the orders O_1 or O_2.*

C. *If there is a direct (inverse) semicycle, $c_1c_2 \cdots c_{2p}$, the element immediately following the semicycle is one of the $2p - 1$ elements succeeding c_{2p}^{-1} in the order $O_3(O_4)$.*

Any finite or infinite sequence of elements satisfying conditions A, B and C will be called *admissible*. Any finite admissible sequence of m elements is the symbol of one, and only one, interval of stage m. Any admissible infinite sequence is the symbol of a point of U.

5. **Additional properties of the Nielsen development.** In order to obtain certain desired theorems, it is necessary to derive additional properties of the Nielsen development. Two of these are of a quantitative nature and will be derived in this paragraph.

LEMMA 5.1. *There exists a positive constant D, depending only on the genus p, such that if P is any point of U, the broken path $k(P)$ (in case there are two, either of the associated paths) of the net n associated with P does not wander a NE distance as great as D from the NE ray OP.*

This is essentially proved by Nielsen (NN pp. 220–222), so the proof will be omitted.

[9] Tsai-Han Kiang, On the groups of orientable two-manifolds, Proceedings of the National Academy of Sciences, vol. 17 (1931), pp. 142–144.

LEMMA 5.2. *Hypothesis: The interval $I = AB$ of U is such that its length, $L(I)$, is not greater than $\pi/4$; A' and B' are diametrically opposite A and B, respectively; L_A and L_A' are circular arcs in Ψ, both passing through A and A' and forming an angle α with U, where $0 < \alpha < \pi/2$; L_B and L_B' are similarly determined with respect to B and B'; R_A is the open region between L_A and L_A'; R_B is the open region between L_B and L_B'; $\bar{R} = R_A \cdot R_B$ is the region common to R_A and R_B.*

Conclusion: There exist positive constants a and $\bar{a}$ determined by α such that if $\underline{d}$ denotes the lower bound of the Euclidean distance from a point of $\bar{R}$ to a point of U, then

$$aL(I) < \underline{d} < \bar{a}L(I) \,. \tag{5.2}$$

For denoting $L(I)$ by 2θ, an elementary computation yields

$$\underline{d} = \frac{1}{ctn\,\alpha + \sin\theta + \sqrt{ctn^2\,\alpha + \sin^2\theta}} \frac{\sin\theta}{\theta} 2\theta\,,$$

and since $0 < \theta < \pi/8$, a and $\bar{a}$ are given by

$$a = \frac{1}{ctn\,\alpha + \sin(\pi/8) + \sqrt{ctn^2\,\alpha + \sin^2(\pi/8)}} \frac{\sin(\pi/8)}{\pi/8}\,,$$

$$\bar{a} = \tfrac{1}{2}\tan\alpha\,.$$

LEMMA 5.3. *If I is an interval of U, and ϕ is a fixed positive angle less than 2π, the locus of points interior to U at which I subtends the angle ϕ through NE rays is an arc of a circle passing through the end points A and B of I. The circle C forms an angle $\phi/2$ with U. If the length, $L(I)$, of I is not greater than $\pi/2$ and $\phi < \pi/2$, there exist positive constants b and $\bar{b}$, determined by ϕ, such that if r denotes the Euclidean radius of C,*

$$bL(I) < r < \bar{b}L(I) \,. \tag{5.3}$$

The proof of this lemma involves only the methods of elementary geometry and will be omitted.

In the remainder of this paragraph, the intervals of U considered will be intervals determined by the process of Nielsen.

THEOREM 5.1. *There exists a positive constant k_1, fixed by the genus p, such that the ratio of any interval to the interval of the preceding stage in which it is contained is greater than the constant k_1.*

Let $I_m = AB$ be the interval $c_1 \cdots c_m$ and I_{m+1} the interval $c_1 \cdots c_m c_{m+1}$. Let Q_{m+1} be the terminal point of the broken path $c_1 \cdots c_m$ and Q_{m+2} the terminal point of the broken path $c_1 \cdots c_m c_{m+1}$. The interval I_{m+1} subtends by NE rays through its end points and Q_{m+1} an angle ϕ_2 at Q_{m+1} such that $\phi_2 \leqq \pi/2p$. If I_{m+2} is the middle interval of I_{m+1} in stage $(m + 2)$, the interval I_{m+2} subtends through NE rays an angle $\phi_3 = \pi/2p$ at Q_{m+2}.

Denoting the NE length of a side of the fundamental polygon S_0 by $\bar{D}$, the points of Ψ which lie at a NE distance $D + \bar{D}$ from the line AA' passing through the origin constitute two circular arcs L_A and L_A', both of which pass

through A and A'. These form an angle α with U such that $0 < \alpha < \pi/2$, α being determined by $D + \bar{D}$, and hence by the genus p. The arcs L_B and L_B' are similarly determined with respect to the line BB'. The regions R_A, R_B and $\bar{R}$ are defined as in Lemma 5.2.

The point Q_{m+1} lies on at least one of the broken paths of n associated with A and also on one of the broken paths associated with B. Hence by Lemma 5.1, its NE distance from either of the lines AA' or BB' is not as great as D. The NE distance from Q_{m+2} to Q_{m+1} is $\bar{D}$, and it follows that Q_{m+2} lies in the region $\bar{R}$. From Lemma 5.2, the minimum Euclidean distance from Q_{m+2} to U must be at least $aL(I_m)$.

Now suppose that Theorem 5.1 were not true. This would mean that given $\epsilon > 0$, intervals I_m and I_{m+1} could be found such that $L(I_{m+1}) < \epsilon L(I_m)$. This would imply $L(I_{m+2}) < \epsilon L(I_m)$. Since the interval I_{m+2} subtends an angle $\pi/2p$ at Q_{m+2}, from Lemma 5.3, Q_{m+2} would lie on a circle C through the end points of I_{m+2} and forming an angle $\pi/4p$ with U. Denoting the Euclidean radius of C by r, from (5.3) would follow

$$r < \bar{b}L(I_{m+2}) < \bar{b}\epsilon L(I_m) . \tag{5.4}$$

Hence the minimum Euclidean distance from Q_{m+2} to U could not exceed $2r < 2\bar{b}\epsilon L(I_m)$. By choosing $\epsilon \lesssim a/2\bar{b}$, a contradiction arises, and Theorem 5.1 holds.

THEOREM 5.2. *The intervals I_1, I_2, I_1' and I_2' of U are given as follows:*

$$\begin{aligned} I_1 &= c_1c_2 \cdots c_md_1 = AB ; & L(I_1) &= \lambda_1 ; \\ I_2 &= c_1c_2 \cdots c_md_1 \cdots d_n = CD ; & L(I_2) &= \lambda_2 ; \\ I_1' &= c_1'c_2' \cdots c_m'd_1 = A'B' ; & L(I_1') &= \lambda_1' ; \\ I_2' &= c_1'c_2' \cdots c_m'd_1 \cdots d_n = C'D' ; & L(I_2') &= \lambda_2' . \end{aligned}$$

The interval I_1 is not an end interval of the interval $c_1 \cdots c_m$ in which it is contained; I_2 is not an end interval of the interval $c_1c_2 \cdots c_md_1 \cdots d_{n-1}$; and similar properties hold for I_1' and I_2'. Then there exist positive constants k_2 and k_3, depending only on the genus p, such that

$$k_2\lambda_1'\lambda_2 < \lambda_1\lambda_2' < k_3\lambda_1'\lambda_2 . \tag{5.5}$$

Let P_1 be the terminal point of c_m in $c_1c_2 \cdots c_md_1$; P_2 that of d_{n-1} in $c_1c_2 \cdots c_md_1 \cdots d_{n-1}d_n$; P_1' that of c_m' in $c_1'c_2' \cdots c_m'd_1$; P_2' that of d_{n-1} in $c_1'c_2' \cdots c_m'd_1 \cdots d_{n-1}d_n$. The circles C_1, C_2, C_1', C_2', with Euclidean radii r_1, r_2, r_1', r_2' are determined by the triples of points (P_1, A, B), (P_1, C, D), (P_1', A', B') and (P_1', C', D'), respectively.

If α_1 is the smallest positive angle between C_1 and U, from Lemma 5.3 it is evident that $\alpha_1 = \pi/4p$. Likewise the smallest positive angle between C_1' and U is $\pi/4p$. Let α_2 be the smallest positive angle between C_2 and U. There is a transformation of the group F which takes the point P_1 into the point P_1'.

This transformation takes the part $d_1 \cdots d_n$ of the broken path $c'_1 \cdots c'_m d_1 \cdots d_n$ into the part $d_1 \cdots d_n$ of the broken path $c_1 \cdots c_m d_1 \cdots d_n$. Since the transformation is conformal, the NE rays through P'_2 which determine $C'D'$, and thus form equal angles $\pi/4p$ with d_n, must be transformed into the NE rays through P_2 which determine the interval CD, or C goes into $C'(D')$ and D into $D'(C')$. The circle C'_2 is thus transformed into the circle C_2, and the smallest positive angle between C'_2 and U must also be α_2.

The following formulas are then readily verified:

$$(5.6)\qquad \begin{cases} \tan \dfrac{\lambda_i}{2} = \dfrac{r_i \sin \alpha_i}{1 - r_i \cos \alpha_i}, & i = 1, 2, \\[2ex] \tan \dfrac{\lambda'_i}{2} = \dfrac{r'_i \sin \alpha_i}{1 - r'_i \cos \alpha_1}, & i = 1, 2. \end{cases}$$

From these there results:

$$(5.7)\qquad \tan \frac{\lambda_1}{2} \tan \frac{\lambda'_2}{2} = \frac{r_1}{r_2}\frac{r'_2}{r'_1}\, \frac{1 - r'_2 \cos \alpha_1}{1 - r'_2 \cos \alpha_2}\, \frac{1 - r_2 \cos \alpha_2}{1 - r_1 \cos \alpha_1} \tan \frac{\lambda_2}{2} \tan \frac{\lambda'_1}{2}.$$

It can be readily shown that the radii r_1, r_2, r'_1 and r'_2 are all uniformly less than a constant less than 1, the constant depending only on the genus, so that the coefficient of $\tan \lambda_2/2 \cdot \tan \lambda'_1/2$ lies between positive constants depending on the genus if this is true of the two ratios r_1/r_2 and r'_1/r'_2. But elementary geometric considerations yield this result. Thus there exist positive constants $\bar{k}_2$ and $\bar{k}_3$ determined by the genus p such that

$$(5.8)\qquad \bar{k}_2 \tan \frac{\lambda'_1}{2} \tan \frac{\lambda_2}{2} < \tan \frac{\lambda_1}{2} \tan \frac{\lambda'_2}{2} < \bar{k}_3 \tan \frac{\lambda'_2}{2} \tan \frac{\lambda_2}{2}.$$

But since λ_1, λ_2, λ'_1 and λ'_2 are all positive and less than $\pi/4$, the inequalities of the theorem follow from (5.8) with $k_2 = \bar{k}_2/4$ and $k_3 = 4\bar{k}_3$.

6. The transformation of the inverse of an admissible symbol into an admissible symbol. By the *inverse* of the symbol

$$(6.1)\qquad c_1 c_2 \cdots c_m,$$

is meant the symbol

$$(6.2)\qquad c_m^{-1} c_{m-1}^{-1} \cdots c_1^{-1}.$$

If (6.1) is admissible, it does not follow that (6.2) is admissible, for though conditions A and B of §4 are satisfied, it is not necessarily true that condition C holds. But it will be shown that there is an admissible symbol having the same number of elements as (6.2) and which is closely related to (6.2) in ways to be described. In particular, considering (6.2) as the geometric configuration made up of a sequence of segments of the net n with the successive sym-

bols of (6.2), there corresponds to this broken line a broken line which is admissible and has the same end points.

It is supposed that (6.2) is not admissible. Then there exists an integer i such that

$$c_i^{-1} \cdots \quad c_1^{-1} \tag{6.3}$$

is not admissible and

$$c_{i-1}^{-1} \cdots \quad c_1^{-1} \tag{6.4}$$

is admissible. The only way in which (6.3) can be inadmissible is in the occurrence of $2p$ elements in order $O_1(O_2)$ such that if the last element in this set of $2p$ elements is c_k^{δ}, the next element in (6.3) is not one of the $(2p - 1)$ elements following $c_k^{-\delta}$ in the order $O_3(O_4)$. Since (6.4) is admissible, it follows that

$$c_i^{-1}c_{i-1}^{-1} \cdots c_{i-(2p-1)}^{-1}c_{i-2p}^{-1} \tag{6.5}$$

is not admissible, and the first $2p$ symbols, $c_i^{-1} \cdots c_{i-(2p-1)}^{-1}$, are either in order O_1 or in order O_2. By the *complement* of such a sequence of $2p$ symbols in order $O_1(O_2)$ will be meant the sequence of $2p$ symbols following c_i in the order $O_2(O_1)$. Geometrically the complement is simply the other half of a polygon of the net n with orientation opposite to that of $c_i^{-1} \cdots c_{i-(2p-1)}^{-1}$. Let the first $2p$ elements in (6.5) be replaced by the complement

$$\bar{c}_i^{-1}\bar{c}_{i-1}^{-1} \cdots \bar{c}_{i-(2p-1)}^{-1}. \tag{6.6}$$

Then the symbol

$$\bar{c}_i^{-1}\bar{c}_{i-1}^{-1} \cdots \bar{c}_{i-(2p-1)}^{-1}c_{i-2p}^{-1} \tag{6.7}$$

is admissible. It may be that

$$\bar{c}_i^{-1}\bar{c}_{i-1}^{-1} \cdots \bar{c}_{i-(2p-1)}^{-1}c_{i-2p}^{-1} \cdots c_1^{-1} \tag{6.8}$$

is admissible. If so, the first step in the desired process is attained.

If not, since the sequence $\bar{c}_{i-(2p-2)}^{-1}\bar{c}_{i-(2p-1)}^{-1}\bar{c}_{i-2p}^{-1}$ is not in either order O_1 or O_2, the first possible inadmissible sequence in (6.8) is

$$\bar{c}_{i-(2p-1)}^{-1}c_{i-2p}^{-1}c_{i-(2p+1)}^{-1} \cdots c_{i-(4p-1)}^{-1}, \tag{6.9}$$

and furthermore, since $c_{i-2p}^{-1} \cdots c_1^{-1}$ is admissible, (6.9) is the only possible inadmissible sequence of $2p + 1$ elements in (6.8). Let the first $2p$ elements in (6.9) be replaced by the complementary set

$$\bar{\bar{c}}_{i-(2p-1)}^{-1}\bar{c}_{i-2p}^{-1}\bar{c}_{i-(2p+1)}^{-1} \cdots \bar{c}_{i-(4p-2)}^{-1}. \tag{6.10}$$

If the symbol

$$\bar{c}_i^{-1}\bar{c}_{i-1}^{-1} \cdots \bar{c}_{i-(2p-2)}^{-1}\bar{\bar{c}}_{i-(2p-1)}^{-1}\bar{c}_{i-2p}^{-1} \cdots \bar{c}_{i-(4p-2)}^{-1}c_{i-(4p-1)}^{-1} \cdots c_1^{-1} \tag{6.11}$$

is admissible, the desired first step is attained. If not, at least the symbol

$$\bar{c}_i^{-1} \cdots \bar{c}_{i-(2p-2)}^{-1}\bar{\bar{c}}_{i-(2p-1)}^{-1}\bar{c}_{i-2p}^{-1} \cdots \bar{c}_{i-(4p-2)}^{-1}c_{i-(4p-1)}^{-1} \tag{6.12}$$

is admissible. By an argument like that already used, that is, by the replacing of a sequence of $2p$ elements by the complementary set, the first element of the replaced sequence being $\bar{c}_{i-(4p-2)}^{-1}$, a greater admissible sequence is obtained. The increase is by $2p - 1$ elements each time, hence eventually, by a finite number of steps, an admissible symbol

$$d_i^{-1}d_{i-1}^{-1} \cdots d_1^{-1} \tag{6.13}$$

corresponding to (6.3) is obtained. This is the desired first step in obtaining the admissible symbol corresponding to (6.2).

This process can be carried through geometrically. Starting with a path CC' having the symbol (6.3) a path DD' having the symbol (6.13) is obtained by a succession of steps, each consisting of replacing $2p$ successive sides of a polygon of the net n by the $2p$ remaining sides of the same polygon. Both aspects will be considered.

We note the following properties of (6.13) as compared with (6.3):

PROP. I. *If c_k^{-1} in* (6.3) *is not one of a sequence of at least* $(2p - 1)$ *elements in either order O_1 or order O_2, $d_k^{-1} = c_k^{-1}$. Geometrically, the segment c_k^{-1} is identical with the segment d_k^{-1} unless the segment c_k^{-1} is one of a sequence of* $(2p - 1)$ *segments of* (6.3) *which are successive segments on a polygon of the net n.*

PROP. II. *The element c_i^{-1} is replaced by either the following or the preceding element in the order O_3. Geometrically, the path* (6.13) *has the same initial point as the path* (6.3).

PROP. III. *The element c_1^{-1} is not changed, $c_1^{-1} = d_1^{-1}$. Geometrically, the last segment of* (6.3) *stays fixed, and hence in particular the terminal point of the path* (6.13) *is the same as the terminal point of the path* (6.3).

PROP. IV. *Geometrically, if c_k^{-1} is a segment of the path* (6.3) *and d_k^{-1} is the corresponding segment of* (6.13), *there are three possibilities:* (1), *the segments are identical;* (2), *d_k^{-1} is a side of the polygon of n on which the sequence defined in Prop. I and containing c_k^{-1} lies; or* (3), *d_k^{-1} is a side of a polygon which has a side in common with the polygon on which the sequence defined in Prop. I and containing c_k^{-1} lies and both c_k^{-1} and d_k^{-1} have end points in common with this common side. In the third case d_k^{-1} can be obtained from c_k^{-1} by the following procedure: c_k^{-1} is first replaced by a segment $\bar{c}_k^{-1}$ having the same terminal point as c_k^{-1} and adjacent to c_k^{-1} in the sequence about its terminal point in the clockwise* (*counterclockwise*) *sense; then this segment is replaced by a segment having the same initial point and adjacent to $\bar{c}_k^{-1}$ at its initial point in the counterclockwise* (*clockwise*) *sense about this initial point.*

From the hypothesis that (6.1) is admissible, it is obvious geometrically that the pair $c_{i+1}^{-1}\, d_i^{-1}$ is not in either order O_1 or order O_2. It follows that if

$$c_m^{-1}\, c_{m-1}^{-1} \cdots c_{i+1}^{-1}\, d_i^{-1} \tag{6.14}$$

is admissible,

$$c_m^{-1}\, c_{m-1}^{-1} \cdots c_{i+1}^{-1}\, d_i^{-1} \cdots d_1^{-1} \tag{6.15}$$

is admissible, and the desired admissible symbol corresponding to (6.2) has been attained.

Assuming that (6.14) is inadmissible, let j be the integer such that

$$c_j^{-1}\, c_{j-1}^{-1} \cdots c_{i+1}^{-1}\, d_i^{-1} \tag{6.16}$$

is inadmissible, but

$$c_{j-1}^{-1} \cdots c_{i+1}^{-1}\, d_i^{-1} \tag{6.17}$$

is admissible. By precisely the process used above, we obtain a symbol

$$d_j^{-1} \cdots d_{i+1}^{-1}\, d_i^{-1} \tag{6.18}$$

corresponding to (6.16) and admissible. By Prop. III d_i^{-1} is retained as the last symbol.

The symbol

$$d_j^{-1} \cdots d_{i+1}^{-1}\, d_i^{-1} \cdots d_1^{-1} \tag{6.19}$$

is admissible. This is evident except in case d_{i+1}^{-1} is identical with the symbol c_i, for then the $2p+1$ elements $d_{i+1}^{-1} \cdots d_{i-(2p-1)}^{-1}$ might be in order O_1 or order O_2. But this is impossible. For this would imply, first, that the $2p$ elements $d_{i+2p}^{-1} \cdots d_{i+1}^{-1}$ are in order $O_1(O_2)$; second, that hence the $2p$ elements $c_{i+(2p-1)}^{-1} \cdots c_{i+1}^{-1}\, c_i^{-1}$ are in the order $O_2(O_1)$; third, since (6.1) is admissible, that c_{i+2p} is one of the $2p-1$ elements following $c_{i+(2p-1)}^{-1}$ in the order $O_4(O_3)$. But Prop. D, relating c_{i+2p}^{-1} to its corresponding segment d_{i+2p}^{-1} cannot be satisfied. Thus (6.19) is admissible.

By a continuation of this process, after a finite number of steps, the admissible symbol

$$d_m^{-1} \cdots d_1^{-1} \tag{6.20}$$

corresponding to (6.2) is attained. Properties I–IV hold for the pair (6.2) and (6.20), provided, of course, that in II, c_i^{-1} is replaced by c_m^{-1}.

THEOREM 6.1. *If $c_1 c_2 \cdots c_m$ is admissible, $d_m^{-1}\, d_{m-1}^{-1} \cdots d_1^{-1}$ is the admissible symbol corresponding to $c_m^{-1} \cdots c_1^{-1}$, and $c_m^{-1} = d_m^{-1}$, then $c_1 \cdots c_m$ is the admissible symbol corresponding to $d_1 \cdots d_m$.*

For if this were not the case, there would be two admissible paths with the same initial point and the same terminal segment. But then there would be two distinct intervals of the same stage with the same symbol, and this is impossible.

7. **The point set E and its invariant properties.** A set $\{g\}$ of directed geodesics on m_R is represented in Ψ by set G of directed geodesics. The set G is invariant under transformations of the group F.

A point of U is determined by a coordinate θ, $0 \leqq \theta < 2\pi$, where θ is the arc-length on U taken in the counterclockwise sense and measured from the point $z = e^{-i\pi/4p}$. An ordered pair P_1P_2 of points on U is then specified by a pair of coordinates (θ_1, θ_2), where θ_1 is the coordinate of P_1 and θ_2 is that of P_2. Thus a directed geodesic, the initial point of which is P_1 and the terminal point P_2 determines a point lying in the rectangle Δ_U, $0 \leqq \theta_i < 2\pi$, $i = 1, 2$, of the (θ_1, θ_2) plane. The set $\{g\}$ of directed geodesics on M_R thus determines a set E in Δ_U.

A transformation T of the group F transforms U into itself, hence $P_1(\theta_1)$ into some $P_1'(\theta_1')$ and $P_2(\theta_2)$ into $P_2'(\theta_2')$. There is thereby defined a transformation T' of the rectangle Δ_U into itself, the point (θ_1', θ_2') corresponding to (θ_1, θ_2). Since the set G is invariant under the group F, the set E is invariant under the group F' of transformations of Δ_U into itself defined by the various transformations of F.

A torus could replace the rectangle Δ_U in the preceding and following considerations.

The set E may or may not be measurable. If it is measurable, it will be shown that it is either of measure zero or of the measure of the rectangle Δ_U. In order to arrive at this result it is necessary to know how the points of E are transformed by the transformations of F'. The technic of Nielsen is extremely useful in this consideration.

As described in §4, U is divided by the process of Nielsen into a net of intervals. Hence each of the intervals $0 \leqq \theta_1 < 2\pi$, $\theta_2 = 0$, and $0 \leqq \theta_2 < 2\pi$, $\theta_1 = 0$, can be divided by the same process into subintervals and these subintervals taken in ordered pairs, one from the first interval and the other from the second, determine a net of rectangles in the rectangle Δ_U. These intervals will be considered open hence the rectangles considered are open sets (the interior points of the rectangles). Since each interval into which U is divided by the process of Nielsen has associated with it a certain finite symbol, $c_1 c_2 \cdots c_m$, each rectangle of the net defined in Δ_U can be denoted by a pair of symbols $\{c_1 \cdots c_m, d_1 \cdots d_n\}$.

Before considering the set E in its entirety, it is necessary to consider the subset in a subrectangle of Δ_U. The particular subrectangle considered will be $\{a_1^2, a_1^{-2}\}$, and it will be denoted by Δ. The subset of E in Δ will be denoted in the usual way by $E \cdot \Delta$.

8. **The measure of the set E** $\cdot\, \Delta$. Assuming that the set E is measurable, a proof is desired that the set $E \cdot \Delta$ is either of measure zero or of the measure of Δ. In order to arrive at this result, several lemmas are required.

LEMMA 8.1. *If δ_1' and δ_2' are subintervals of $\delta_1 = a_1^2$ and $\delta_2 = a_1^{-2}$, respectively, and T is a transformation of F such that $\bar\delta_1 = T(\delta_1')$ and $\bar\delta_2 = T(\delta_2')$ are subintervals of δ_1 and δ_2 respectively, there exist positive constants h_1 and h_2 such that*

$$h_1 < \frac{L(\delta_1')\, L(\delta_2')}{L(\bar\delta_1)\, L(\bar\delta_2)} < h_2 . \tag{8.1}$$

Let z_1 and z_1' be the end points of δ_1', z_2 and z_2' the end points of δ_2', $\bar{z}_1$ and $\bar{\bar{z}}_1$ the end points of $\bar{\delta}_1$, and $\bar{z}_2$, $\bar{\bar{z}}_2$ the end points of $\bar{\delta}_2$. Let T be given by

$$w = \frac{\alpha z + \bar{\beta}}{\beta z + \bar{\alpha}}, \qquad \alpha\bar{\alpha} - \beta\bar{\beta} = 1\,.$$

A brief computation shows that

$$(8.2) \qquad \frac{|z_1 - z_1'| \cdot |z_2 - z_2'|}{|\bar{z}_1 - \bar{\bar{z}}_1| \cdot |\bar{z}_2 - \bar{\bar{z}}_2|} = \frac{|z_1 - z_2| \cdot |z_1' - z_2'|}{|\bar{z}_1 - \bar{z}_2| \cdot |\bar{\bar{z}}_1 - \bar{\bar{z}}_2|}\,.$$

Since δ_1 and δ_2 are intervals of U with intervals separating them, the right side of (8.2) is bounded above and away from zero. Since the intervals δ_1 and δ_2 are not greater than semicircles, the ratio of the chord joining two points of $\delta_1(\delta_2)$ to the arc joining these points lies between two positive constants. The left side of (8.2) is that obtained from the ratio in (8.1) by replacing arcs by the corresponding chords, and the statement of the lemma follows.

Let Δ' be the rectangle of Δ_U determined by the intervals δ_1' and δ_2' of Lemma 8.1. As described in §7, the transformation T defines a transformation of Δ' into $\bar{\Delta}$, the rectangle of Δ_U determined by $\bar{\delta}_1$ and $\bar{\delta}_2$. In particular, if T is a transformation of the group F, the corresponding transformation T' of Δ_U into itself leaves the set E invariant, hence the set $E' = E \cdot \Delta'$ is transformed into the set $\bar{E} = E \cdot \bar{\Delta}$. Assuming that the set E is measurable, each of the sets E' and $\bar{E}$ is measurable, and the following lemma holds.

LEMMA 8.2. $h_1 \mu(\bar{E}) \leqq \mu(E') \leqq h_2 \mu(\bar{E})$.

For since the set E' is measurable, it can be enclosed in a denumerable set of rectangles S_R' in Δ', the sum of whose areas differs from $\mu(E')$ by less than an arbitrary positive constant ϵ. The transformation T' takes the set S_R' into a set $\bar{S}_R$ of rectangles in $\bar{\Delta}$, and the set $\bar{S}_R$ encloses the set $\bar{E}$. From Lemma 8.1, the sum of the areas of the rectangles of the set $\bar{S}_R$i s less than $h_1^{-1}[\mu(E') + \epsilon]$. Passing to the limit, the measure of the set $\bar{E}$ cannot be greater than $\mu(E')/h_1$. By a similar argument it can be shown that if $\mu(\bar{E}) < \mu(E')/h_2$ a contradiction arises, and the desired lemma follows.

It will now be assumed that the set $E \cdot \Delta$ is of positive measure, and it will be shown[10] that the measure of $E \cdot \Delta$ must then be equal to the measure or area of the rectangle Δ.

Let

$$(8.3) \qquad \{a_1^2 c_1 \cdots c_\nu s_1 s_2 \cdots s_\lambda e_1 a_1 a_1^2,\ a_1^{-2} d_1 \cdots d_\nu d_{\nu+1}\}$$

be a rectangle of the net in Δ_U subject to the following conditions:

COND. 1. *The interval* $a_1^2 c_1 \cdots c_\nu s_1$ *is the middle interval of the interval* $a_1^2 c_1 \cdots c_\nu$.

[10] W. Seidel, in a paper, Note on a metrically transitive system, Proceedings of the National Academy of Sciences, April, 1933, p. 453, considered a problem which has much in common with the problem considered here, but it has not appeared evident that his methods apply readily to the present situation.

COND. 2. *The interval* $a_1^2 c_1 \cdots c_\nu s_1 s_2$ *is the middle interval of the interval* $a_1^2 c_1 \cdots c_\nu s_1$.

COND. 3. *The admissible symbol corresponding to the symbol* $a_1^{-1} e_1^{-1} s_\lambda^{-1} \cdots s_1^{-1}$ *is of the form* $a_1^{-2} \bar{s}_\lambda^{-1} \cdots \bar{s}_2^{-1} s_1^{-1}$.

COND. 4. *The interval* $a_1^{-2} d_1 \cdots d_\nu d_{\nu+1}$ *is the middle interval of* $a_1^{-2} d_1 \cdots d_\nu$.

The existence of rectangles satisfying conditions 1, 2 and 4 is evident. That cond. 3 can be satisfied is shown by the following property.

PROP. A′. *If* $a_1^2 c_1 \cdots c_\nu s_1 s_2 \cdots s_{\lambda-2}$ *is admissible,* $s_{\lambda-1}, s_\lambda$ *and* e_1 *can be chosen so that condition* 3 *is satisfied.*

For let $s_{\lambda-1}$ be chosen so that $a_1^2 c_1 \cdots c_\nu s_1 \cdots s_{\lambda-2} s_{\lambda-1}$ is the middle interval of $a_1^2 c_1 \cdots c_\nu s_1 \cdots s_{\lambda-2}$. Then there are $4p - 1$ $(p > 1)$ possible choices for s_λ, and s_λ can be chosen from among these such that: (a), it does not follow $s_{\lambda-1}$ in either order O_1 or order O_2; (b), it is not a_1^{-1}; and (c), it does not precede a_1 in either order O_1 or order O_2. Then e_1 can be chosen as a_1. The admissible symbol corresponding to $a_1^{-2} s_\lambda^{-1} s_{\lambda-1}^{-1} \cdots s_1^{-1}$ is then of the form $a_1^{-2} s_\lambda^{-1} s_{\lambda-1}^{-1} \bar{s}_{\lambda-2}^{-1} \cdots \bar{s}_3^{-1} \bar{s}_2^{-1} s_1^{-1}$, from Prop. I of §6.

Furthermore, the following important property holds.

PROP. B′. *If* $a_1^{-2} t_\lambda \cdots t_5$ *is admissible,* $s_3, \cdots s_\lambda, e_1$ *can be so chosen that conditions* 1, 2 *and* 3 *are satisfied and* $\bar{s}_\lambda^{-1} = t_\lambda, \cdots \bar{s}_5^{-1} = t_5$.

For let s_4 be chosen such that $a_1^{-2} t_\lambda \cdots t_5 s_4^{-1}$ is admissible and t_5 and s_4^{-1} are not in either order O_1 or O_2. Let s_3 be chosen such that $a_1^2 c_1 \cdots c_\nu s_1 s_2 s_3$ is admissible and such that $a_1^{-2} t_\lambda \cdots t_5 s_4^{-1} s_3^{-1}$ is admissible with the last two elements not in order O_1 or O_2. The admissible symbol corresponding to $s_3 s_4 t_5^{-1} \cdots t_\lambda^{-1} a_1^2$ will be of the form $s_3 s_4 \bar{t}_5^{-1} \cdots \bar{t}_\lambda^{-1} \bar{e}_1 a_1$. Let $s_i = \bar{t}_i^{-1}$, $i = 5, \cdots \lambda$, and $e_1 = \bar{e}_1$. The symbol $a_1^2 c_1 \cdots c_\nu s_1 \cdots s_\lambda e_1 a_1$ is then admissible, and from Theorem 6.1, the admissible symbol corresponding to $a_1^{-1} e_1^{-1} s_\lambda^{-1} \cdots s_1^{-1}$ will be of the form $a_1^{-2} t_\lambda \cdots t_5 s_4^{-1} s_3^{-1} s_2^{-1} s_1^{-1}$. This proves Prop. B′.

The admissible symbol corresponding to the inverse of $a_1^2 c_1 \cdots c_\nu s_1 s_2 \cdots s_\lambda e_1 a_1$, subject to conditions 1 to 3, will be of the form $a_1^{-2} \bar{s}_\lambda^{-1} \cdots \bar{s}_3^{-1} \bar{s}_2^{-1} s_1^{-1} \bar{c}_\nu^{-1} \cdots \bar{c}_1^{-1} e_2 a_1^{-1}$, where $c_1 \cdots c_\nu$ completely determines $\bar{c}_\nu^{-1} \cdots \bar{c}_1^{-1} e_2$.

If, in determining the rectangle (8.3), $c_1 \cdots c_\nu$ and $d_1 \cdots d_{\nu+1}$ and λ are kept fixed, but $s_1 s_2 \cdots s_\lambda$ is subject only to the stated conditions, a set σ_λ of rectangles is obtained, all of which lie in the rectangle $\{a_1^2 c_1 \cdots c_\nu, a_1^{-2} d_1 \cdots d_\nu\}$. Let this last rectangle be denoted by Δ_1. It will be supposed in the following that $\lambda > 6$. The area of the set σ_λ will be denoted by the usual symbol for measure, $\mu(\sigma_\lambda)$.

LEMMA 8.3. *There exists a positive constant* k_4, *independent of* Δ_1 *and* λ, *such that*

$$\mu(\sigma_\lambda) > k_4 \mu(\Delta_1) .$$

Using the notation of §5,

$$\mu(\Delta_1) = L(a_1^2 c_1 \cdots c_\nu) \cdot L(a_1^{-2} d_1 \cdots d_\nu) .$$

The following inequalities follow at once from Theorem 5.1.

$$L(a_1^{-2}\, d_1 \cdots d_\nu\, d_{\nu+1}) > k_1\, L(a_1^{-2}\, d_1 \cdots d_\nu)\,.$$

$$L(a_1^2\, c_1 \cdots c_\nu\, s_1\, s_2) > k_1^2\, L(a_1^2\, c_1 \cdots c_\nu)\,.$$

$L(a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_\lambda\, e_1\, a_1\, a_1^2) > k_1^6\, L(a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_{\lambda-2})$. From Prop. A′, $s_3 \cdots s_{\lambda-2}$ is subject only to the condition that $a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_{\lambda-2}$ be admissible. It follows that if $\sum_i L_i$ denotes the sum of the lengths of the sides of the set σ_λ which are parallel to the θ_1-axis,

$$\sum_i L_i > k_1^6\, L(a_1^2\, c_1 \cdots c_\nu\, s_1\, s_2)\,.$$

Combining the various inequalities,

$$\mu(\sigma_\lambda) = L(a_1^{-2}\, d_1 \cdots d_{\nu+1}) \sum_i L_i > k_1^9\, L(a_1^{-2} d_1 \cdots d_\nu)\cdot L(a_1^2\, c_1 \cdots c_\nu)\,,$$

or

$$\mu(\sigma_\lambda) > k_1^9\, \mu(\Delta_1)\,.$$

Setting $k_4 = k_1^9$, Lemma 8.3 follows.

The rectangle $\Delta' = \{a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_\lambda\, e_1\, a_1\, a_1^2,\ a_1^{-2}\, d_1 \cdots d_\nu\, d_{\nu+1}\}$ is determined by the two intervals $\delta_1' = a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_\lambda\, e_1\, a_1\, a_1^2$ and $\delta_2' = a_1^{-2}\, d_1 \cdots d_\nu\, d_{\nu+1}$. Each of these intervals has associated with it a directed broken geodesic, the origin being the vertex in each case, as described in §4. Let P_1 be the terminal point of the broken geodesic $a_1^2\, c_1 \cdots c_\nu\, s_1 \cdots s_\lambda\, e_1\, a_1$. There is one and only one transformation of the group F taking P_1 into the origin. This transformation transforms the interval δ_1' into the interval $\delta_1 = a_1^2$. The interval δ_2' is carried into an interval $\bar\delta_2$ determined by the following procedure: from O the sequence $a_1^{-1}\, e_1^{-1}\, s_\lambda^{-1} \cdots s_1^{-1}\, c_\nu^{-1} \cdots c_1^{-1}\, a_1^{-2}\, a_1^{-2}\, d_1 \cdots d_{\nu+1}$ of geodesic segments of the net n is traced; from the initial point of the last segment NE rays forming equal angles of magnitude $\pi/4p$ with the last segment are drawn; these rays intersect U in the end points of $\bar\delta_2$. The symbol of $\bar\delta_2$ is then $a_1^{-2}\bar s_\lambda^{-1} \cdots \bar s_2^{-1} s_1^{-1} \bar c_\nu^{-1} \cdots \bar c_1^{-1} e_2 a_1^{-1} a_1^{-2} d_1 \cdots d_{\nu+1}$, where $a_1^{-2}\bar s_\lambda^{-1} \cdots \bar s_2^{-1} s_1^{-1}$ is the admissible symbol corresponding to $a_1 e_1^{-1} s_\lambda^{-1} \cdots s_1^{-1}$ and $s_1^{-1} \bar c_\nu^{-1} \cdots \bar c_1^{-1} e_2 a_1^{-1}$ is the admissible symbol corresponding to $s_1^{-1} c_\nu^{-1} \cdots c_1^{-1} a_1^{-2}$. Hence there is a transformation of the group F' taking the rectangle Δ' into the rectangle $\bar\Delta = \{a_1^2,\ a_1^{-2}\bar s_\lambda^{-1} \cdots \bar s_2^{-1} s_1^{-1} \bar c_\nu^{-1} \cdots \bar c_1^{-1} e_2 a_1^{-1} a_1^{-2} d_1 \cdots d_{\nu+1}\}$, which is also a subrectangle of Δ.

From Lemma 8.1,

$$h_1\, \mu(\bar\Delta) < \mu(\Delta') < h_2\, \mu(\bar\Delta)\,.$$

From Lemma 8.2

$$h_1\, \mu(E\cdot\bar\Delta) \leqq \mu(E\cdot\Delta') \leqq h_2\, \mu(E\cdot\bar\Delta)\,,$$

where E is the invariant set of §7.

To each of the rectangles of the set σ_λ a transformation of the group F' can be applied such that the resulting rectangle is of the form $\bar{\Delta}$. Let the resulting set of rectangles for fixed λ be denoted by $\bar{\sigma}_\lambda$.

LEMMA 8.4. *There exists a positive constant k_5, determined by the genus p, such that if λ is chosen sufficiently large,*

$$\mu(E\cdot\bar{\sigma}_\lambda) \geqq k_5\,\mu(\bar{\sigma}_\lambda)\cdot\mu(E\cdot\Delta)\,.$$

A rectangle of the set $\bar{\sigma}_\lambda$ is of the form

$$\{a_1^2,\ a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_2^{-1}s_1^{-1}\bar{c}_\nu^{-1}\cdots\bar{c}_1^{-1}e_2a_1^{-1}a_1^{-2}d_1\cdots d_\nu d_{\nu+1}\}\,,$$

where, from Prop. B$'$, the symbol $\bar{s}_\lambda^{-1}\cdots\bar{s}_5^{-1}$ is subject only to the condition that $a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_5^{-1}$ be admissible.

The lemma will first be proved in the special case that $E\cdot\Delta$ is a rectangle $\{a_1^2f_1\cdots f_q,\ a_1^{-2}g_1\cdots g_r\}$ of the net in Δ. If $\lambda > r+4$, there are then rectangles of the set $\bar{\sigma}_\lambda$ which intersect the rectangle $E\cdot\Delta$. Let ${}_\lambda I_j$, $j = 1, 2, \cdots J$, denote the set of intervals $a_1^{-2}\bar{s}_\lambda^{-1}\cdots d_{\nu+1}$. Then

$$\mu(\bar{\sigma}_\lambda) = L(a_1^2)\sum_{j=1}^{J}L({}_\lambda I_j)\,.$$

Let ${}_\lambda I_k$, $k = j_1, \cdots j_K$, denote the intervals of the set ${}_\lambda I_j$, $j = 1, \cdots J$, which are in $a_1^{-2}g_1\cdots g_r$. Then

$$\mu(E\cdot\bar{\sigma}_\lambda) = L(a_1^2f_1\cdots f_q)\sum_{k=j_1}^{j_K}L({}_\lambda I_k)\,.$$

For each $\lambda > r+4$, let ${}_\lambda I_2'$ be some fixed one of the possible intervals ${}_\lambda I_j$, $j = 1, \cdots J$, e.g., ${}_\lambda I_2' = a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_2^{-1}s_1^{-1}\bar{c}_\nu^{-1}\cdots d_{\nu+1}$. Let ${}_\lambda I_1'$ be the interval $a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_3^{-1}\bar{s}_2^{-1}$, and if ${}_\lambda I_j = a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_2^{-1}s_1^{-1}c_\nu^{-1}\cdots d_{\nu+1}$, let ${}_\lambda\bar{I}_j = a_1^{-2}\bar{s}_\lambda^{-1}\cdots\bar{s}_2^{-1}$. Applying Theorem 5.2,

$$k_2L({}_\lambda I_1')\,L({}_\lambda I_j) < L({}_\lambda\bar{I}_j)\,L({}_\lambda I_2') < k_3L({}_\lambda I_1')\,L({}_\lambda I_j)\,, \tag{8.4}$$

where k_2 and k_3 are positive constants determined by the genus. Summing with respect to j, λ fixed,

$$k_2L({}_\lambda I_1')\sum_{j=1}^{J}L({}_\lambda I_j) < L({}_\lambda I_2')\sum_{j=1}^{J}L({}_\lambda\bar{I}_j) < L({}_\lambda I_2')\,L(a_1^{-2})\,. \tag{8.5}$$

The relation (8.4) holds in particular with ${}_\lambda I_k$ in place of ${}_\lambda I_j$, and summing, there results

$$\sum_{k=j_1}^{j_K}L({}_\lambda I_k) > k_3^{-1}\frac{L({}_\lambda I_2')}{L({}_\lambda I_1')}\sum_{k=j_1}^{j_K}L({}_\lambda\bar{I}_k)\,.$$

In the set ${}_\lambda\bar{I}_k$, $k = j_1, \cdots j_K$, occur all possible symbols of the form $a_1^{-2} g_1 \cdots g_r \bar{s}_{\lambda-r+1}^{-1} \cdots \bar{s}_5^{-1}$, hence with the aid of Theorem 5.1,

$$\sum_{k=j_1}^{j_K} L({}_\lambda \bar{I}_k) > k_1^3 \, L(a_1^{-2} g_1 \cdots g_r) \, ,$$

and thus

$$\sum_{k=j_1}^{j_K} L({}_\lambda I_k) > k_1^3 \, k_3^{-1} \frac{L({}_\lambda I_2')}{L({}_\lambda I_1')} \, L(a_1^{-2} g_1 \cdots g_r) \, . \tag{8.6}$$

Combining inequalities (8.5) and (8.6)

$$\mu(E \cdot \bar{\sigma}_\lambda) = L(a_1^2 f_1 \cdots f_q) \sum_{k=j_1}^{j_K} L({}_\lambda I_k) > k_1^3 \, k_2 \, k_3^{-1} \, [L(a_1^{-2})]^{-2} \, \mu(E \cdot \Delta) \, \mu(\bar{\sigma}_\lambda) \, .$$

Setting

$$k_5 = 2^{-1} \, k_1^3 \, k_2 \, k_3^{-1} \, [L(a_1^{-2})]^{-2},$$

the lemma follows in the special case under consideration.

It is evident that the lemma holds in the case that $E \cdot \Delta$ is the sum of a finite number of non-overlapping rectangles of the net in Δ.

If $E \cdot \Delta = \sum_{i=1}^{\infty} R_i$, a sum of non-overlapping rectangles of the net in Δ, given $\epsilon > 0$, there exists an N such that $\mu(\sum_{i=1}^{N} R_i) > (1 - \epsilon) \, \mu(E \cdot \Delta)$. For this finite set of rectangles the lemma holds, with $2k_5$ in place of k_5, and hence for λ sufficiently large,

$$\mu(E \cdot \bar{\sigma}_\lambda) \geqq \mu\left(\bar{\sigma}_\lambda \cdot \sum_{i=1}^{N} R_i\right) > 2k_5 \, \mu(\bar{\sigma}_\lambda) \, \mu\left(\sum_{i=1}^{N} R_i\right) > 2k_5 (1 - \epsilon) \, \mu(\bar{\sigma}_\lambda) \, \mu(E \cdot \Delta) \, .$$

It is only necessary to choose $\epsilon < 1/4$ and the lemma holds with $3k_5/2$ in place of k_5.

If $E \cdot \Delta$ is an open set, it is the sum of a denumerable set of non-overlapping rectangles of the net, together with a set of measure zero, since the rectangles are all assumed open. The lemma again holds with the factor 3/2 on the right.

Finally, since $E \cdot \Delta$ is measurable, given $\epsilon > 0$, there exists an open set E_0 in Δ such that $E_0 \supset E \cdot \Delta$ and $\mu(E \cdot \Delta) > \mu(E_0) - \epsilon$. It readily follows that

$$\epsilon > \mu(E_0 \cdot \bar{\sigma}_\lambda) - \mu(E \cdot \bar{\sigma}_\lambda) \geqq 0 \, .$$

From the preceding, for sufficiently large λ,

$$\mu(E_o \cdot \bar{\sigma}_\lambda) \geqq \tfrac{3}{2} \, k_5 \, \mu(\bar{\sigma}_\lambda) \, \mu(E_0) \, . \tag{8.7}$$

From Lemma 8.2, $\mu(\bar{\sigma}_\lambda) > h_2^{-1} \mu(\sigma_\lambda)$, and from Lemma 8.3, $\mu(\sigma_\lambda) > k_4 \, \mu(\Delta_1)$, hence $\mu(\bar{\sigma}_\lambda)$ is bounded away from zero. The same is evidently true of $\mu(E_0)$, of

$\mu(E_0 \cdot \bar{\sigma}_\lambda)$ from (8.7) and hence of $\mu(E \cdot \bar{\sigma}_\lambda)$. But since ϵ can be chosen arbitrarily small, (8.7) yields

$$\mu(E \cdot \bar{\sigma}_\lambda) \geqq k_5 \, \mu(\bar{\sigma}_\lambda) \, \mu(E \cdot \Delta) , \tag{8.8}$$

for λ sufficiently large.

The fundamental theorem can now be obtained.

THEOREM 8.1. *If the set of $E \cdot \Delta$ is of positive measure it is of the measure of Δ.*

For the following inequalities hold for λ sufficiently large,

$$\mu(E \cdot \sigma_\lambda) \geqq h_1 \, \mu(E \cdot \bar{\sigma}_\lambda) ,$$

$$\mu(E \cdot \bar{\sigma}_\lambda) \geqq k_5 \, \mu(\bar{\sigma}_\lambda) \, \mu(E \cdot \Delta) ,$$

$$\mu(\bar{\sigma}_\lambda) > h_2^{-1} \, \mu(\sigma_\lambda) ,$$

$$\mu(\sigma_\lambda) > k_4 \, \mu(\Delta_1) .$$

Combining these with the obvious one

$$\mu(E \cdot \Delta_1) \geqq \mu(E \cdot \sigma_\lambda) ,$$

there results

$$\mu(E \cdot \Delta_1) \geqq k_6 \, \mu(\Delta_1) , \tag{8.9}$$

where $k_6 = h_1 h_2^{-1} k_4 k_5 \, \mu(E \cdot \Delta) > 0$.

If $E \cdot \Delta$ were not of the measure of Δ, there would be an interior point P of Δ at which the metric density of the set E would exist and be zero. Thus for a sufficiently small square Δ_2 with center at P, the following inequality would hold

$$\mu(E \cdot \Delta_2) < \tfrac{1}{2} k_6 \, \mu(\Delta_2) , \tag{8.10}$$

with Δ_2 lying in Δ. But a denumerable set of non-overlapping rectangles of the form Δ_1 can be chosen such that their sum is equal to Δ_2 except for a set of measure zero. For each of these rectangles (8.9) holds, and hence (8.10) cannot hold. This proves Theorem 8.1.

9. **Transitive geodesics on M_R.** A directed geodesic, g, on M_R is represented in Ψ by an infinite set G of directed geodesics and these in turn determine a point set E_G of Δ_U. The directed geodesic g is *transitive* if the point set E_G is everywhere dense in Δ_U. This is equivalent to the property that, δ and δ' being arbitrary intervals of U, there is a member of the set G with initial point in δ and terminal point in δ'. It is obvious that this definition is equivalent to the one usually given in terms of the phase-space.

If g is not transitive, it is *intransitive*. A member of the set G is *transitive* or *intransitive* according as g is transitive or intransitive. If g is transitive (intransitive) any point of the set E_G is *transitive* (*intransitive*).

The points of U can be divided into two mutually exclusive classes, F_T and

F_I, of transitive and intransitive points, respectively. A point $P(\theta)$ belongs to F_T, or is *transitive*, if all the geodesics in Ψ with one end point at P are transitive. P belongs to F_I if this is not the case. It is known[11] that the set F_T is of the measure of the length of U, or 2π.

Let the set of transitive points of Δ_U be denoted by E_T.

THEOREM 9.1. $\mu(E_T) = 4\pi^2 = \mu(\Delta_U)$.

For the set E_I of intransitive points of Δ_U is such that the point $Q(\theta_1, \theta_2)$ is a point of E_I if, and only if, both of the points $P(\theta_1)$ and $P(\theta_2)$ of U are intransitive. The projection of the set E_I on each of the axes is then a set of measure zero and it follows that the set E_I is itself of measure zero. The statement of the theorem is equivalent to this.

With this result it is now easy to extend Theorem 8.1 to the whole set E.

THEOREM 9.2. *If E is the set of Δ_U determined by a set of directed geodesics on M_R, and E is measurable, either $\mu(E) = 0$, or $\mu(E) = \mu(\Delta_U) = 4\pi^2$.*

From Theorem 8.1, either $\mu(E\cdot\Delta) = 0$ or $\mu(E\cdot\Delta) = \mu(\Delta)$. It will be sufficient to consider the second case, as the first can be reduced to the second by considering the set complementary to E in Δ_U. Let $\Delta - E\cdot\Delta$ denote the points of Δ not in E. Then $\mu(\Delta - E\cdot\Delta) = 0$. Each of the transformations T_i, $i = 1, 2, \cdots$, of F determines a transformation T_i', $i = 1, 2, \cdots$, of F'. It is easily verified that each of the transformations T_i', $i = 1, 2, \cdots$, is one-to-one and analytic with non-vanishing Jacobian except on two line segments parallel, respectively, to the coordinate axes. (There would be no exceptions if a torus were used in place of Δ_U). Hence each of the sets $T_i'(\Delta - E\cdot\Delta)$ is a set of measure zero, and $\mu[E + \sum_{i=1}^{\infty} T_i'(\Delta - E\cdot\Delta)] = \mu(E)$. The set $E + \sum_{i=1}^{\infty} T_i'(\Delta - E\cdot\Delta)$ includes all the points of Δ, and, like E, it is invariant under any of the transformations T_i', $i = 1, 2, \cdots$. But if P_T is a transitive point of Δ_U, the set $T_i'(P_T)$, $i = 1, 2, \cdots$, is everywhere dense in Δ_U, and there must be one of the set in Δ. Thus the set $\sum_{i=1}^{\infty} T_i'(\Delta)$, which is included in $E + \sum_{i=1}^{\infty} T_i'(\Delta - E\cdot\Delta)$ includes all the transitive points and from Theorem 9.1 it follows that $\mu(E) = \mu(E + \sum_{i=1}^{\infty} T_i'(\Delta - E\cdot\Delta)) = \mu(\Delta_U) = 4\pi^2$.

10. **Measure of sets on M_R and the set H.** If to the interior points of the fundamental region S_0 is added a properly chosen set of $2p$ sides and just one vertex, a set of points in one-to-one correspondence with the points of M_R is obtained. It is understood in the following that S_0 stands for this enlarged region.

A set H' of points in S_0 will be said to be measurable if it is measurable in the sense of Lebesgue, considering the (x, y) plane as a Euclidean plane. The measure of the set H' is defined as

$$4\int_{H'} (1 - x^2 - y^2)^{-2}\, dA\,, \tag{10.1}$$

where the integral is the L-integral over the set H'.

[11] Myrberg, loc. cit., Hedlund, loc. cit.

Given a set H on M_R, measurability and measure will be defined as for the set H' of S_0 corresponding to H. To prove that a set H is of measure zero it is evidently sufficient to prove that the corresponding set H' is of measure zero considering the (x, y) plane as a Euclidean plane.

Given a set $\{g\}$ of directed geodesics on M_R, the remaining directed geodesics on M_R will be denoted by $C\{g\}$. Let $\bar{H}$ be the set of points of M_R, at any one of which, P, the measure of the directions of the members of $\{g\}$ passing through P is 2π. The set similarly defined for $C\{g\}$ will be denoted by $\bar{\bar{H}}$.

THEOREM 10.1. *If the sets E and $\bar{H}$ are measurable, $\bar{H}$ is either of measure zero or of the measure of the whole of M_R. In the first case, the measure of the set $\bar{\bar{H}}$ is that of M_R, and in the second, $\bar{\bar{H}}$ is of measure zero.*

From Theorem 9.2, the set E is either of measure zero or of the measure of Δ_U. The second case will be considered first. The measure of the set $\bar{H}$ is determined by that of the set $\bar{H}'$ of S_0 corresponding to it. Let the angular coordinate, θ, at each point of S_0 be measured in the positive sense from the geodesic ray drawn from the point to the point $z = e^{-i\pi/4p}$.

Let γ be a NE straight line which intersects S_0 in a segment. If s is the (Euclidean) arc-length on γ measured from one of the end points on U, let the segment of γ in S_0 be $s_1 < s < s_2$. Then the pair (s, θ) determines a directed geodesic in Ψ, and hence a point (θ_1, θ_2) of Δ_U. The pair (s, θ) will determine a directed geodesic corresponding to one in the set $\{g\}$ provided that the point (θ_1, θ_2) belongs to the set E of Δ_U corresponding to $\{g\}$. The transformation from (s, θ) to (θ_1, θ_2) is analytic with non-vanishing Jacobian in the rectangle

$$s_1 < s < s_2, \qquad 0 < \theta < 2\pi, \tag{10.2}$$

with the exception of the points for which $\theta = \pi$, and the points of two analytic arcs, where continuity does not hold. Hence the points of (10.2) which determine geodesics in the set $\{g\}$ will be a set of the measure of the whole rectangle. From a theorem of Fubini,[12] the points of the segment, $s_1 < s < s_2$, of γ, which belong to the set $\bar{H}'$ will be of linear measure equal to that of the segment. But the NE straight line γ is arbitrary, hence on every NE straight line segment in S_0, the points of $\bar{H}'$ will be of the measure of the whole segment, and by an obvious generalization of the theorem of Fubini, the set $\bar{H}'$, being measurable, is of the measure of S_0 itself. The set $\bar{H}$ is thus of the measure of M_R.

At any point of the set $\bar{\bar{H}}$ the measure of the directions of the set $\{g\}$ passing through this point is zero. But then the set $\bar{\bar{H}}$ is a subset of the set of M_R complementary to $\bar{H}$, and since $\bar{H}$ is of measure equal to that of M_R, the set $\bar{\bar{H}}$ is of measure zero.

The remainder of the theorem is easily obtained by interchanging the rôles of $\{g\}$ and $C\{g\}$.

11. The phase-space and metrical transitivity. Let P be a point of M_R and P' the corresponding point of S_0. A direction in M_R at P is determined by

[12] Carathéodory, Vorlesungen über Reelle Funktionen, p. 627.

an angular coordinate, θ, at P', where θ will be measured as in §10. A point P and a direction θ at P is then determined by the ordered pair (P, θ), and the totality of such ordered pairs, or elements, will be the *phase-space* M.

The points of M are in one-to-one correspondence with the set M', (x, y, θ), where (x, y) is a point of S_0 and θ is the θ of M. A set E_M of points of M will be measurable if the set $E_{M'}$ of M' corresponding to E_M is Lebesgue measurable. The measure of a measurable set $E_{M'}$ will be defined as

$$4 \int_{E_{M'}} (1 - x^2 - y^2)^{-2}\, dv\,, \tag{11.1}$$

where the integral is the L-integral. The measure of the set E_M of M will be defined as the measure of the corresponding set $E_{M'}$ of M'. To prove that the set M' is of measure zero it is sufficient to prove that it is of measure zero, considering (x, y, θ) as Cartesian coordinates in a Euclidean three-space.

The totality of elements on a set $\{g\}$ of directed geodesics on M_R determines a set E_M of M. The set E will again be the two-dimensional set of §7 determined by the set $\{g\}$.

THEOREM 11.1. *If the set E_M is a measurable set, the set E is measurable, and conversely.*

A point (x, y, θ) of M determines a point $P(x, y)$ of S_0 and a direction at this point, hence a directed geodesic $\bar{g}$ in Ψ. Let the initial point of $\bar{g}$ be $P_1(\theta_1)$, the terminal point $P_2(\theta_2)$ and s the (Euclidean) arc-length of the segment P_1P of $\bar{g}$. The transformation, $\overline{T}$, from (x, y, θ) to (θ_1, θ_2, s) is analytic with non-vanishing Jacobian for (x, y) in Ψ and $0 < \theta < 2\pi$, with the exception of two analytic surfaces where the continuity does not hold. Let M_Δ be the set of points (x, y, θ) of M' such that the directed geodesic of Ψ determined by a point of M_Δ has its initial point in $\delta_1 = a_1^2$ and its terminal point in $\delta_2 = a_1^{-2}$.

A directed geodesic g of M_R determines a set of points of M' consisting of a set of analytic curve segments, the number being finite only if g is periodic. Hence the set $E_{M'} \cdot M_\Delta$, which is a measurable set if the set E_M is measurable, will either be empty or consist of a set of analytic arcs, each of these arcs corresponding to a geodesic in Ψ with initial point, $P_1(\theta_1)$, in δ_1, and terminal point, $P_2(\theta_2)$, in δ_2. The points (θ_1, θ_2) thus determined belong to the set $E \cdot \Delta$ of §7. By transforming from the coordinates (x, y, θ) to (θ_1, θ_2, s), these arcs in M' are transformed into straight lines parallel to the s-axis, and the measurability of the set $E \cdot \Delta$ is implied by the measurability of the set $E_{M'}$. From Theorem 8.1, the set $E \cdot \Delta$ is either of measure zero or of the measure of Δ, for though it was assumed in the proof of Theorem 8.1 that the set E was measurable, the proof required only the measurability of the set $E \cdot \Delta$. It is easily seen that the proof of Theorem 9.2 can be modified so as to require only the measurability of the set $E \cdot \Delta$. Hence the measurability of the set E_M implies that the set E is either of measure zero or of the measure of Δ_U.

Conversely, if the set E is measurable, it is known that it is either of measure zero or else this is true of the complement in Δ_U. Let $s(\theta_1, \theta_2)$ be the Euclidean

arc-length of the geodesic with initial point $P_1(\theta_1)$ and terminal point $P_2(\theta_2)$. If E is of measure zero, it is evident that the set (θ_1, θ_2, s), where (θ_1, θ_2) belongs to E and $0 < s < s(\theta_1, \theta_2)$, is of measure zero. But the set $\overline{T}(E_{M'})$ is a subset of the last set, and the set $E_{M'}$ must be of measure zero. Then E_M is measurable and of measure zero. The other case, when E is not of measure zero, reduces to the last by consideration of complements.

Denoting the complement of E_M in M by $C(E_M)$, the following theorem is contained in the proof of Theorem 11.1.

THEOREM 11.2. *If the set E_M is measurable, either $\mu(E_M) = 0$ or $\mu[C(E_M)] = 0$.*

This is the statement of the metrical transitivity of the geodesics on M_R.

It was assumed in §10 that the set $\bar{H}$ was measurable. But this is implied by the measurability of the set E. For, from Theorem 11.2, if E is measurable, either $\mu(E_M) = 0$ or $\mu[C(E_M)] = 0$. In the first case, $\mu[C(E_M)] = \mu(M)$ and from the theorem of Fubini the set $\bar{H}$ is of measure equal to that of Δ_U. This implies $\mu(\bar{H}) = 0$. In the second case, $\mu(E_M) = \mu(M)$ and by the same theorem $\mu(\bar{H}) = \mu(M_R)$.

THEOREM 11.3. *The measurability of the set E implies the measurability of the set $\bar{H}$.*

12. **On the metrical transitivity of sets on U.** A simple corollary of the preceding results is that invariant measurable sets on U must have a property of metrical transitivity embodied in the following theorem.

THEOREM 12.1. *If a set, E_U, of points on U, is measurable and invariant under the transformations of the group F, it is either of measure zero or of the measure of U.*

For let $E_U \times E_U$ be the set of points of Δ_U such that if $P(\theta_1, \theta_2)$ is any point of $E_U \times E_U$, the points $P_1(\theta_1)$ and $P_2(\theta_2)$ are both members of E_U. The set $E_U \times E_U$ is a measurable set and $\mu(E_U \times E_U) = [\mu(E_U)]^2$. The set $E_U \times E_U$ is invariant under the transformations of the group F' and from Theorem 9.2, either $\mu(E_U \times E_U) = 0$ or $\mu(E_U \times E_U) = 4\pi^2 = [\mu(U)]^2$. The statement of the theorem follows at once.

13. **Remarks.** The manifold considered here is restricted in at least two senses. The group F considered is a single group among the totality of Fuchsian groups which could be used to define closed orientable manifolds of genus greater than one. It seems likely that the technic used here can be readily extended to include these other groups.

The manifold is restricted to being of constant negative curvature. It seems likely that the result holds when the curvature is negative but not necessarily constant. Whether the methods used in the case of constant curvature extend to the more general case does not seem obvious.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

Annals of Mathematics
Vol 35, No. 4, October, 1934

TOPOLOGICAL IMMERSION OF PEANIAN CONTINUA IN A SPHERICAL SURFACE[1]

By Schieffelin Claytor

(Received September 26, 1933)

Introduction

In 1930 Kuratowski[2] established the following result:

Theorem A: *A peanian continuum,*[3] *containing but a finite number of simple closed curves, is homeomorphic with a subset of the plane, provided that it does not contain a topological image of either of the complexes* $\mathfrak{C}$ *and* $\mathfrak{D}$, *where*

$\mathfrak{C}$ *consists of two groups of three vertices each and nine* 1*-cells, in a fashion that each vertex of one group together with each vertex of the other group bounds in a* 1*-cell;*

$\mathfrak{D}$ *consists of five vertices and ten* 1*-cells, in a fashion that each pair of vertices bounds a* 1*-cell.*[4]

This theorem suggests the more general problem treated in this paper, namely, the characterization of the peanian continua which are homeomorphic with a subset of the surface of a sphere.

We find it convenient, first, to designate as a *primitive skew curve* any point-set homeomorphic with either of the complexes $\mathfrak{C}$ and $\mathfrak{D}$; and then, to distinguish among the set of all peanian continua a class, $\mathfrak{R}$, consisting of those which fail to contain a *primitive skew curve.*

Certain boundary sets are defined in an arbitrary continuum, $\mathbf{M}$, of class $\mathfrak{R}$, as follows:

Definition I: *A simple closed curve,* $\mathbf{T}$, *in* $\mathbf{M}$ *is called a boundary curve of* $\mathbf{M}$ *provided that there do not exist in* $\mathbf{M}\text{-}\mathbf{T}$ *distinct components* E *and* F *such that*

(α) *a point-pair of* $\mathbf{F}(E)$[5] *separates a point pair of* $\mathbf{F}(F)$ *on* $\mathbf{T}$, *or*

(β) $\mathbf{F}(E) \equiv \mathbf{F}(F) = a + b + c$; *where* a, b, c, *are distinct points.*

---

[1] Presented to the American Mathematical Society, April 14, 1933.

[2] Sur les problème des courbes gauches en Topologie, Fundamenta Mathematicae, vol. 15, pp. 271–283.

[3] Peanian continuum = a continuous image of a linear interval = a *compact, connected,* and *locally connected* continuum.

[4] A combinatorial equivalent of Theorem A has been given by H. Whitney; Non-separable and Planar Graphs, Trans. Amer. Soc., vol. 34 (1932), pp. 339–362. Also, Mazurkiewicz has recently shown that if M is a one-dimensional peanian continuum which can through no "kleine Abbildungen" be carried into a plane curve, then M necessarily contains a topological image of either $\mathfrak{C}$ or $\mathfrak{D}$. Über nicht plattbare Kurven, Fund. Math., vol. 20, pp. 281–284.

[5] For $L \subset \mathbf{M}$, $\mathbf{F}(L) = \overline{L} \cdot (\overline{M \dot{-} L})$ = the boundary of L in $\mathbf{M}$.

DEFINITION II: *A simple continuous arc, B, in* **M** *is called a boundary arc of* **M** *provided that if* **C** *is a cyclic element of* **M** *containing more than one point of B, then the point-set $B \cdot$* **C** *is a subset of some boundary curve of* **C**. *A point of* **M** *is called a boundary point of* **M** *provided that it lies on some boundary arc of* **M**.

We summarize our results:

(A) *If a continuum of class $\mathfrak{N}$ is without a boundary curve, then its proper*[6] *cyclic elements are simple closed surfaces.*

(B) *Every continuum of class $\mathfrak{N}$ is homeomorphic with a subset of some peanian continuum of which the proper cyclic elements are simple closed surfaces.*

(C) *A necessary and sufficient condition that a continuum,* **M**, *of class $\mathfrak{N}$ be homeomorphic with a subset of a spherical surface is that each cut-point, P, of* **M** *be a boundary point of the closure of every component of* **M** $-$ *P.*

Corollaries:

(a) *A simple closed surface may be characterized as a cyclic continuum of class $\mathfrak{N}$, other than a point, which does not have a boundary curve.*

(b) *Every cyclic continuum of class $\mathfrak{N}$ is homeomorphic with a subset of a spherical surface.*

(c) *A closed 2-cell may be characterized as a cyclic continuum of class $\mathfrak{N}$ that has precisely one boundary curve.*

The result (C) furnishes a complete solution of our central problem, for it is clear that a peanian continuum cannot fail to belong to the class $\mathfrak{N}$ and at the same time have a topological image in the surface of a sphere.

The author takes this opportunity to express his grateful appreciation of the encouragement and counsel that he has received of Professor John R. Kline in the solution of the problems of this paper.

I. Preliminary Lemmas

Throughout our work, **M** will be used to denote an arbitrary continuum of class $\mathfrak{N}$; **C**, to denote an arbitrary cyclic continuum of class $\mathfrak{N}$ that is distinct from a point; and **T** to denote any simple closed curve in **M** (or in **C**). A *region* is an *open* and *connected* point-set.

LEMMA 1. *A necessary and sufficient condition that* **T** *fail to be a boundary curve of* **M** *is that either*

(a) *arcs ab and cd exist in* **M** *having no point in common and only their end points on* **T** *such that $(a + b)$ separates $(c + d)$ on* **T**; *or if not, then*

(b) *arc triods*[7] *T_1 and T_2 exist in* **M** *such that $T_1 \cdot T_2 = T_1 \cdot \mathbf{T} = T_2 \cdot \mathbf{T} = a + b + c$, where a, b, c are the feet of the triods T_1 and T_2.*

---

[6] A cyclic element of a peanian continuum is said to be *proper* if it consists of more than a single point.

[7] An *arc triod* is a point-set consisting of three simple continuous arcs that have a common end point P which is the only point common to any pair of the arcs. The point P is called the *center* of the triod, while its (three) end points are called its *feet*.

Demonstration. The necessity of our condition is quite obvious, so we consider its sufficiency. There are two cases to be dealt with:

Case 1. Assume that condition (a) holds. Then the segments $\widehat{ab}$[8] and $\widehat{cd}$ do not both lie in a single component D of $\mathbf{M} - \mathbf{T}$; for otherwise, on taking in D an arc ef that has only the end point e on $\widehat{ab}$ and only the end point f on $\widehat{cd}$, it is readily seen that the point-set $\mathbf{T} + ab + cd + ef$ is a *primitive skew curve* (homeomorphic with the complex $\mathfrak{C}$); and this contradicts our assumption that $\mathbf{M}$ is a continuum of class $\mathfrak{R}$. Accordingly, let E and F be the distinct components of $\mathbf{M} - \mathbf{T}$ that contain, respectively, the sets $\widehat{ab}$ and $\widehat{cd}$. It is then clear that the condition (α) of Definition I holds with respect to E and F.

Case 2. Assume that condition (b) holds, while condition (a) *fails* to hold. The treatment of this case varies but slightly from that of the preceding one.

Corollary. *If* $\mathbf{T}$ *is a boundary curve of* $\mathbf{M}$ *and* L *is a peanian subcontinuum of* $\mathbf{M}$ *that contains* $\mathbf{T}$, *then* $\mathbf{T}$ *is also a boundary curve of* L.

Definitions. *Distinct components* E *and* F *of* $\mathbf{C} - \mathbf{T}$ *are said to be on opposite sides of* $\mathbf{T}$ *provided* $\mathbf{T}$ *is not a boundary curve of the continuum* $\mathbf{T} + E + F$.

If $E_1, E_2, \cdots, E_n$, $(n > 1)$, *is a finite collection of distinct components of* $\mathbf{C} - \mathbf{T}$ *such that* E_i *and* E_{i+1} *are on opposite sides of* $\mathbf{T}$, $1 \leqq i \leqq n - 1$, *then the point-set* $E_1 + E_2 + \cdots + E_n$ *is called a chain in* $\mathbf{C} - \mathbf{T}$ *joining* E_1 *to* E_n, *and is denoted by* $C[E_1, E_n]$. *The order of a chain is the number of distinct components contained in it.*

A nest, $n(E)$, *determined in* $\mathbf{C} - \mathbf{T}$ *by one of its components* E, *is the point-set consisting of* E *itself together with all the components* E_x *of* $\mathbf{C} - \mathbf{T}$ *for which there exists a chain joining* E *and* E_x.

A few of the properties of *nests* in $\mathbf{C} - \mathbf{T}$ are as follows:

Property 1. *If* F *is a component of a nest* $n(E)$, *then*

$$n(E) \equiv n(F).$$

Property 2. *For an arbitrary pair of nests* $n(E)$ *and* $n(F)$ *we have*

$$n(E) \equiv n(F) \qquad or \qquad n(E) \cdot n(F) = 0.$$

The demonstrations of Properties 1 and 2 are left to the reader, and we proceed with

Property 3. *The chains joining two arbitrary components of a nest in* $\mathbf{C} - \mathbf{T}$ *are either all of even order, or else all of odd order.*

Demonstration. On supposing our proposition false, it readily follows that there exist, in $\mathbf{C} - \mathbf{T}$, *chains* $C_1[E, F]$ and $C_2[E, F]$ which are of *even* and *odd* order respectively, and which, furthermore, *have in common only the components* E *and* F. For convenience we let $E_1, E_2, \cdots, E_{2n}$ represent uniquely the components of $C_1[E, F]$ and $E_{2n}, E_{2n+1}, \cdots, E_{2m}$ represent uniquely the components of $C_2[E, F]$, in such a way that E_i and E_{i+1} are on opposite sides of $\mathbf{T}$, for $i = 1, 2, \cdots, 2m - 1$, while $E_1 = E_{2m} = E$ and $E_{2n} = F$.

[8] If ab is a simple continuous arc then $\widehat{ab}$ denotes the point-set $ab - (a + b)$.

DEFINITION II: *A simple continuous arc, B, in* **M** *is called a boundary arc of* **M** *provided that if* **C** *is a cyclic element of* **M** *containing more than one point of B, then the point-set $B \cdot$* **C** *is a subset of some boundary curve of* **C**. *A point of* **M** *is called a boundary point of* **M** *provided that it lies on some boundary arc of* **M**.

We summarize our results:

(A) *If a continuum of class $\mathfrak{R}$ is without a boundary curve, then its proper*[6] *cyclic elements are simple closed surfaces.*

(B) *Every continuum of class $\mathfrak{R}$ is homeomorphic with a subset of some peanian continuum of which the proper cyclic elements are simple closed surfaces.*

(C) *A necessary and sufficient condition that a continuum,* **M**, *of class $\mathfrak{R}$ be homeomorphic with a subset of a spherical surface is that each cut-point, P, of* **M** *be a boundary point of the closure of every component of* **M** $-$ *P.*

Corollaries:

(a) *A simple closed surface may be characterized as a cyclic continuum of class $\mathfrak{R}$, other than a point, which does not have a boundary curve.*

(b) *Every cyclic continuum of class $\mathfrak{R}$ is homeomorphic with a subset of a spherical surface.*

(c) *A closed 2-cell may be characterized as a cyclic continuum of class $\mathfrak{R}$ that has precisely one boundary curve.*

The result (C) furnishes a complete solution of our central problem, for it is clear that a peanian continuum cannot fail to belong to the class $\mathfrak{R}$ and at the same time have a topological image in the surface of a sphere.

The author takes this opportunity to express his grateful appreciation of the encouragement and counsel that he has received of Professor John R. Kline in the solution of the problems of this paper.

I. Preliminary Lemmas

Throughout our work, **M** will be used to denote an arbitrary continuum of class $\mathfrak{R}$; **C**, to denote an arbitrary cyclic continuum of class $\mathfrak{R}$ that is distinct from a point; and **T** to denote any simple closed curve in **M** (or in **C**). A *region* is an *open* and *connected* point-set.

LEMMA 1. *A necessary and sufficient condition that* **T** *fail to be a boundary curve of* **M** *is that either*

(a) *arcs ab and cd exist in* **M** *having no point in common and only their end points on* **T** *such that $(a + b)$ separates $(c + d)$ on* **T**; *or if not, then*

(b) *arc triods*[7] *T_1 and T_2 exist in* **M** *such that $T_1 \cdot T_2 = T_1 \cdot \mathbf{T} = T_2 \cdot \mathbf{T} = a + b + c$, where a, b, c are the feet of the triods T_1 and T_2.*

---

[6] A cyclic element of a peanian continuum is said to be *proper* if it consists of more than a single point.

[7] An *arc triod* is a point-set consisting of three simple continuous arcs that have a common end point P which is the only point common to any pair of the arcs. The point P is called the *center* of the triod, while its (three) end points are called its *feet*.

Demonstration. The necessity of our condition is quite obvious, so we consider its sufficiency. There are two cases to be dealt with:

Case 1. Assume that condition (a) holds. Then the segments $\widehat{ab}$[8] and $\widehat{cd}$ do not both lie in a single component D of $\mathbf{M} - \mathbf{T}$; for otherwise, on taking in D an arc ef that has only the end point e on $\widehat{ab}$ and only the end point f on $\widehat{cd}$, it is readily seen that the point-set $\mathbf{T} + ab + cd + ef$ is a *primitive skew curve* (homeomorphic with the complex $\mathfrak{C}$); and this contradicts our assumption that $\mathbf{M}$ is a continuum of class $\mathfrak{R}$. Accordingly, let E and F be the distinct components of $\mathbf{M} - \mathbf{T}$ that contain, respectively, the sets $\widehat{ab}$ and $\widehat{cd}$. It is then clear that the condition (α) of Definition I holds with respect to E and F.

Case 2. Assume that condition (b) holds, while condition (a) *fails* to hold. The treatment of this case varies but slightly from that of the preceding one.

Corollary. *If* $\mathbf{T}$ *is a boundary curve of* $\mathbf{M}$ *and* L *is a peanian subcontinuum of* $\mathbf{M}$ *that contains* $\mathbf{T}$, *then* $\mathbf{T}$ *is also a boundary curve of* L.

Definitions. *Distinct components* E *and* F *of* $\mathbf{C} - \mathbf{T}$ *are said to be on opposite sides of* $\mathbf{T}$ *provided* $\mathbf{T}$ *is not a boundary curve of the continuum* $\mathbf{T} + E + F$.

If $E_1, E_2, \cdots, E_n$, $(n > 1)$, *is a finite collection of distinct components of* $\mathbf{C} - \mathbf{T}$ *such that* E_i *and* E_{i+1} *are on opposite sides of* $\mathbf{T}$, $1 \leqq i \leqq n - 1$, *then the point-set* $E_1 + E_2 + \cdots + E_n$ *is called a chain in* $\mathbf{C} - \mathbf{T}$ *joining* E_1 *to* E_n, *and is denoted by* $C[E_1, E_n]$. *The order of a chain is the number of distinct components contained in it.*

A nest, $n(E)$, *determined in* $\mathbf{C} - \mathbf{T}$ *by one of its components* E, *is the point-set consisting of* E *itself together with all the components* E_x *of* $\mathbf{C} - \mathbf{T}$ *for which there exists a chain joining* E *and* E_x.

A few of the properties of *nests* in $\mathbf{C} - \mathbf{T}$ are as follows:

Property 1. *If* F *is a component of a nest* $n(E)$, *then*

$$n(E) \equiv n(F)\,.$$

Property 2. *For an arbitrary pair of nests* $n(E)$ *and* $n(F)$ *we have*

$$n(E) \equiv n(F) \qquad \textit{or} \qquad n(E) \cdot n(F) = 0\,.$$

The demonstrations of Properties 1 and 2 are left to the reader, and we proceed with

Property 3. *The chains joining two arbitrary components of a nest in* $\mathbf{C} - \mathbf{T}$ *are either all of even order, or else all of odd order.*

Demonstration. On supposing our proposition false, it readily follows that there exist, in $\mathbf{C} - \mathbf{T}$, *chains* $C_1[E, F]$ and $C_2[E, F]$ which are of *even* and *odd* order respectively, and which, furthermore, *have in common only the components* E *and* F. For convenience we let $E_1, E_2, \cdots, E_{2n}$ represent uniquely the components of $C_1[E, F]$ and $E_{2n}, E_{2n+1}, \cdots, E_{2m}$ represent uniquely the components of $C_2[E, F]$, in such a way that E_i and E_{i+1} are on opposite sides of $\mathbf{T}$, for $i = 1, 2, \cdots, 2m - 1$, while $E_1 = E_{2m} = E$ and $E_{2n} = F$.

[8] If ab is a simple continuous arc then $\widehat{ab}$ denotes the point-set $ab - (a + b)$.

It is now a simple matter to choose subsets $e_1, e_2, \cdots, e_{2m}$ of $E_1, E_2, \cdots, E_{2m}$, respectively, so that

(1) $e_1 \equiv e_{2m}$;

(2) e_i is *closed*[9] relative to E_i for $i = 1, 2, \cdots, 2m$;

(3) The closure $\bar{e}_i$ of e_i is, for each value of i, either

i) an arc having only its end points on $\mathbf{T}$, or else

ii) an *arc* triod having only its *feet* on $\mathbf{T}$;

(4) e_i and e_{i+1} are on opposite sides of $\mathbf{T}$ for $i = 1, 2, \cdots, 2m - 1$.

Then, on setting $\Psi = \mathbf{T} + e_1 + e_2 + \cdots + e_{2m}$, we see by (2) and (3) that Ψ is a *peanian* continuum that contains *at most a finite* number of simple closed curves. Furthermore, Ψ is without doubt a continuum of class $\mathfrak{R}$. Hence, by Theorem A,[10] there exist a topological image Ψ' of Ψ in a spherical surface S. Denote by $\mathbf{T}', e_1', e_2', \cdots, e_{2m}'$ the images in Ψ' that correspond respectively to $\mathbf{T}, e_1, e_2, \cdots, e_{2m}$ in Ψ. Then (1) implies

(5) $e_1' \equiv e_{2m}'$, and from (4) it follows that

(6) e_i' and e_{i+1}' are on opposite sides of $\mathbf{T}$ for $i = 1, 2, \cdots, 2m - 1$.

By referring to (6), and the corollary to Lemma 1, it is seen that the components e_i' and e_{i+1}', $(1 \leqq i \leqq 2m - 1)$, of $\Psi' - \mathbf{T}'$ *do not both lie in the same domain complementary to* $\mathbf{T}'$ *in* S. Consequently, on properly designating the two components of $S - \mathbf{T}'$ by I_1 and I_2, we obtain

(7) $I_1 \supset e_1' + e_3' + \cdots + e_{2m-1}' \supset e_1'$, and

(8) $I_2 \supset e_2' + e_4' + \cdots + e_{2m}' \supset e_{2m}'$.

It is now evident that (5), (7), and (8) imply a contradiction, for manifestly the *single* point-set represented by e_1' and e_{2m}' cannot belong to *both* I_1 and I_2. We conclude then that our proposition is valid.

LEMMA 2. *At most a finite number of the components of* $\mathbf{C} - \mathbf{T}$ *are of diameter greater than a given positive quantity.*

Demonstration. Assume, on the contrary, that there exists an infinite sequence

(1) $E_1, E_2, \cdots, E_n, \cdots$ of distinct components E_i of $\mathbf{C} - \mathbf{T}$ for which

(2) $\delta(E_i) > r > 0, \qquad i = 1, 2, \cdots, n, \cdots.$

No generality is lost in supposing that this component sequence converges to a sequential limit L in $\mathbf{C}$; then from (2) we get $\delta(L) > 0$. The continuum L is a subset of the simple closed curve $\mathbf{T}$, and hence is an arc; for, if $u(E_i, \mathbf{T})$ denotes the *upper distance*[11] of E_i from $\mathbf{T}$, then according to a lemma due to Zippin[12] we have

---

[9] A subset A of a point-set E is said to be "*closed*" relative to E provided that the limit points of A which belong to E also belong to A.

[10] See Introduction.

[11] The *upper distance* of a point-set M from a point-set N is the upper bound of the numbers $\rho(m, N)$, where m is a point of M.

[12] L. Zippin, "A study of continuous curves and their relation to the Janiszewski-Mullikan Theorem," Trans. Amer. Math. Soc., vol. 31 (1929), pp. 744–770. See lemma of paragraph 3.

(3) $$\lim_{i=\infty} u(E_i, \mathbf{T}) = 0.$$

Now take six distinct points $a_0, a_1, a_2, a_3, a_4, a_5$ that lie on L, and hence on $\mathbf{T}$, in the order named. Then since $\mathbf{C}$ is locally connected we may choose about $a_0, a_1, \cdots, a_5$ respectivley, regions $U_0, U_1, \cdots, U_5$ of $\mathbf{C}$ such that

(i) $\overline{U}_i \cdot \overline{U}_j = 0$, when $i \neq j$; and
(ii) if x_i is an arbitrary point of $\mathbf{T}$ in U_i, $(0 \leqq i \leqq 5)$, then the point order $x_0 x_1 x_2 x_3 x_4 x_5$ is preserved on $\mathbf{T}$.

By (3), there exists an integer $\bar{n}$ that insures

(4) $$E_n \cdot U_i \neq 0, i = 0, 1, \cdots, 5, \text{ for } n > \bar{n},$$

it being assumed that E_n is an element of the sequence (1). Accordingly, let E_{n0}, E_{n1}, E_{n2} be distinct components of $\mathbf{C} - \mathbf{T}$, appearing in (1), for which

(5) n_0, n_1, n_2 are each $> \bar{n}$; and then put

(6) $W_\sigma = U_\sigma + E_{n\sigma} + U_{\sigma+3}, \sigma = 0, 1, 2.$

Since the U's and E's are *connected open* subsets of $\mathbf{C}$ it follows from (4), (5), and (6) that W_σ is also a *connected open* subset of $\mathbf{C}$. Then from

$$W_\sigma \supset a_\sigma + a_{\sigma+3}, \sigma = 0, 1, 2,$$

it is clear that, for each value of σ, W_σ contains an arc $a_\sigma a_{\sigma+3}$ joining the point a_σ of U_σ to the point $a_{\sigma+3}$ of $U_{\sigma+3}$. In the light of the relation (i) it is apparent that the arc $a_\sigma a_{\sigma+3}$ has a *last* point b_σ in the point-set $\mathbf{T} \cdot \overline{U}_\sigma$, and *following* b_σ a *first* point $b_{\sigma+3}$ in the point-set $\mathbf{T} \cdot \overline{U}_{\sigma+3}$. Let $b_\sigma b_{\sigma+3}$ denote the subarc on $a_\sigma a_{\sigma+3}$ that joins b_σ to $b_{\sigma+3}$; and set

$$\Pi = \mathbf{T} + b_0 b_3 + b_1 b_4 + b_2 b_5.$$

With the aid of (ii), it is now easily seen that Π is a primitive skew curve homeomorphic with the complex $\mathfrak{C}$. However, $\mathbf{C} \supset \Pi$. Hence we have obtained the contradiction that $\mathbf{C}$ is not a continuum of class $\mathfrak{N}$; consequently, our lemma is proved.

THE NORMAL REPRESENTATION OF $\mathbf{C}$. *If* $\mathbf{C} - \mathbf{T}$ *is resolved into the sum of two of its subsets A and B in such a way that*

(1) *A and B are separated in* $\mathbf{C}$ *by* $\mathbf{T}$, *and*

(2) $\mathbf{T}$ *is a boundary curve of both* $A + \mathbf{T}$ *and* $\mathbf{T} + B$;

then the exhibition of $\mathbf{C}$ *as the sum* $A + \mathbf{T} + B$ *is called a normal representation of* $\mathbf{C}$.

LEMMA 3. *For each simple closed curve* $\mathbf{T}$ *of* $\mathbf{C}$ *there exists a normal representation:* $\mathbf{C} = A + \mathbf{T} + B$.

Demonstration. Consider an arbitrary nest $n(E_\sigma)$ that is determined in

$\mathbf{C} - \mathbf{T}$ by one of its components E_σ. We define a normal representation of the continuum $\mathbf{T} + n(E_\sigma)$ as follows:

(I) Let A_σ consist of E_σ together with those components E_x of $n(E_\sigma)$ that are joined to E_σ by a *chain* in $\mathbf{C} - \mathbf{T}$ of *odd* order.

(II) Let B_σ consist of those components E_y of $n(E_\sigma)$ that are joined to E_σ by a *chain* in $\mathbf{C} - \mathbf{T}$ of *even* order. It is now clear that either of the assumptions

(1) a component of $n(E_\sigma)$ belongs to both A_σ and B_σ,

(2) $\mathbf{T}$ is not a *boundary curve* of both $A_\sigma + \mathbf{T}$ and $\mathbf{T} + B_\sigma$,

will lead directly to a contradiction of the Property 3 relating to nests in $\mathbf{C} - \mathbf{T}$. Consequently we have the normal representation

$$\mathbf{T} + n(E_\sigma) = A_\sigma + \mathbf{T} + B_\sigma . \tag{3}$$

As a result of Lemma 2 it follows that the aggregate of components in $\mathbf{C} - \mathbf{T}$ is at most countably infinite; hence, the set of distinct nests in $\mathbf{C} - \mathbf{T}$ is countable and may be arranged in a simple sequence

$$n(E_1),\ n(E_2),\ \cdots,\ n(E_\sigma),\ \cdots . \tag{4}$$

Now evidently

$$\mathbf{C} = \mathbf{T} + \sum n(E_\sigma) = \sum [\mathbf{T} + n(E_\sigma)] ;$$

then on referring to relation (3) we get

$$\mathbf{C} = \sum (A_\sigma + \mathbf{T} + B_\sigma) . \tag{5}$$

Finally, set

$$A = A_1 + A_2 + \cdots + A_\sigma + \cdots ,$$

$$B = B_1 + B_2 + \cdots + B_\sigma + \cdots ;$$

and substitute in equation (5) to get

$$\mathbf{C} = A + \mathbf{T} + B ,$$

the desired normal representation.

Lemma 4. *Let* $\{A_i + \mathbf{T}_i + B_i\}$ *be an infinite sequence of normal representations of* $\mathbf{C}$ *such that*

(1) *the sequence* $\{\mathbf{T}_i\}$ *is equicontinuous*[13] *and converges to a simple closed curve* $\mathbf{T}$,

---

[13] A collection G of simple closed curves in a metric space is said to be *equicontinuous* provided that for a given positive number ϵ there exists a positive number δ such that if P and Q are points on a curve $\mathfrak{T}$ of G for which $\rho(P, Q) < \delta$, then there is an arc PQ on $\mathfrak{T}$ of diameter $< \epsilon$. Cf. R. L. Moore, *Concerning certain equicontinuous systems of curves*, Trans. Amer. Math. Soc., vol. 22 (1921), p. 42; and by the same author: *On the generation of a simple surface . . . .*, Fund. Math., T. IV, p. 106.

(2) *the sequences* $\{A_i + \mathbf{T}_i\}$ *and* $\{\mathbf{T}_i + B_i\}$ *converge respectively to sequential limits* $A + \mathbf{T}$ *and* $\mathbf{T} + B$, *where neither* A *nor* B *contains a point of* $\mathbf{T}$;

then $A + \mathbf{T} + B$ *is a normal representation of* $\mathbf{C}$.

Demonstration. It is quite evident that $\mathbf{C} - \mathbf{T}$ is the sum of the two point-sets A and B which are *separated* in $\mathbf{C}$ by $\mathbf{T}$. Furthermore, $\mathbf{T}$ is a *boundary curve* of both $A + \mathbf{T}$ and $\mathbf{T} + B$; for on assuming the contrary, we can obtain, without difficulty, the contradiction that $\mathbf{T}_i$ is *not* a *boundary curve* of both $A_i + \mathbf{T}_i$ and $\mathbf{T}_i + B_i$ for some sufficiently large integer i. The details of this discussion are left to the reader, and we proceed with

LEMMA 5. *If* $A_0 + \mathbf{T}_0 + B_0$ *is a normal representation of* $\mathbf{C}$, *and* $\mathbf{T}'$ *a simple closed curve in* $\mathbf{T}_0 + B_0$, *then there exists a normal representation* $\mathbf{C} = A' + \mathbf{T}' + B'$ for which $A' \supset A_0$ and $B' \subset B_0$.

Demonstration. Excluding the trivial case in which $\mathbf{T}' \equiv \mathbf{T}_0$, we assume that $\mathbf{T}' \not\equiv \mathbf{T}_0$, and uniquely arrange the components a_i of $\mathbf{T}' - \mathbf{T}_0$ in a sequence

$$a_1, a_2, \cdots, a_n, \cdots, \tag{1}$$

that may be either finite or infinite. It is now easily shown that to each element a_n of (1) there may be associated a component b_n of $\mathbf{T}_0 - \mathbf{T}'$ such that

(2) $\bar{a}_n$ separates $\mathbf{T}_0 + \mathbf{B}_0$ between b_n and $\mathbf{T}' - \bar{a}_n$ when $\mathbf{T}' - \bar{a}_n \neq 0$; and

(3) $b_i \neq b_j$ when $i \neq j$.

This particular choice of b_n insures that, on setting

$$\mathbf{T}_n = (\mathbf{T}_{n-1} - b_n) + a_n,$$

the point-sets

$$\mathbf{T}_1, \mathbf{T}_2, \cdots, \mathbf{T}_n, \cdots \tag{4}$$

are each simple closed curves. It is furthermore evident that either $\mathbf{T}_n = \mathbf{T}'$ for some value of $n \geqq 1$, or else that $\lim_{n=\infty} \mathbf{T}_n = \mathbf{T}'$, in which case $\{\mathbf{T}_n\}$ is an equicontinuous collection of curves.

With reference to the curve $\mathbf{T}_1$ we now indicate a way of choosing a normal representation

$$\mathbf{C} = A_1 + \mathbf{T}_1 + B_1 \text{ for which } A_1 \supset A_0 \text{ and } B_1 \subset B_0. \tag{5}$$

To this end, consider the nests in $\mathbf{C} - \mathbf{T}_1$. As distinct nests have no points in common (Property 2), there is a unique nest $n(E_\sigma)$ that contains b_1, where the defining component E_σ of $n(E_\sigma)$ is chosen as *that* particular one which contains b_1 (Property 1). Strictly according to the method of Lemma 3, we now effect a normal representation

$$\mathbf{T}_1 + n(E_\sigma) = A_\sigma + \mathbf{T}_1 + B_\sigma \tag{6}$$

with the result that

$$A_\sigma \supset E_\sigma \tag{7}$$

and furthermore that

$$(8) \qquad n(E_\sigma)\cdot A_0 = A_\sigma .$$

Let us turn next to the nests $n(E)$ in $\mathbf{C} - \mathbf{T}_1$ that are distinct from $n(E_\sigma)$. These fall into two groups $\sum n(E_\xi)$ and $\sum n(E_\delta)$ (either of which may be vacuous) according as

$$(9) \qquad \mathbf{F}\,[n(E)] \subset \mathbf{T}_1 - a_1$$

or

$$(10) \qquad \mathbf{F}\,[n(E)] \subset \bar{a}_1 .$$

A nest $n(E_\xi)$ of $\sum n(E_\xi)$, satisfying relation (9), coincides precisely with a nest in $\mathbf{C} - \mathbf{T}_0$. For each of these set

$$(11) \qquad A_\xi = A_0\cdot n(E_\xi), \qquad B_\xi = B_0\cdot n(E_\xi) ,$$

and obtain the normal representation

$$(12) \qquad \mathbf{T}_1 + n(E_\xi) = A_\xi + \mathbf{T}_1 + B_\xi .$$

A nest $n(E_\delta)$ of $\sum n(E_\delta)$, satisfying relation (10), has no point in A_0, that is,

$$(13) \qquad A_0\cdot n(E_\delta) = 0 .$$

For each of these take an arbitrary normal representation

$$(14) \qquad \mathbf{T}_1 + n(E_\delta) = A_\delta + \mathbf{T}_1 + B_\delta .$$

Relations (12) and (14) yield

$$(15) \qquad \mathbf{T}_1 + \sum n(E_\xi) = \sum A_\xi + \mathbf{T}_1 + \sum B_\xi$$

and

$$(16) \qquad \mathbf{T}_1 + \sum n(E_\delta) = \sum A_\delta + \mathbf{T}_1 + \sum B_\delta .$$

Now set

$$A_1 = A_\sigma + \sum A_\xi + \sum A_\delta , \qquad B_1 = B_\sigma + \sum B_\xi + \sum B_\delta ,$$

and combine (6), (15), and (16) to obtain the normal representation (5), where $A_1 \supset A_0$ and $B_1 \subset B_0$ because of (7), (8), (11), and (13).

For the curve $\mathbf{T}_2$ in (4), the condition (2) now aids in establishing $\mathbf{T}_2 \subset \mathbf{T}_1 + B_1$; hence, by proceeding as we did in getting (5), we can construct a normal representation $\mathbf{C} = A_2 + \mathbf{T}_2 + B_2$, where $A_2 \supset A_1$, $B_2 \subset B_1$; furthermore, $\mathbf{T}_3 \subset \mathbf{T}_2 + B_2$.

On continuing in this way indefinitely, we might possibly reach an integer n for which $\mathbf{T}_n = \mathbf{T}'$; then, on putting $A_n = A'$ and $B_n = B'$, it follows that the normal representation $\mathbf{C} = A' + \mathbf{T}' + B'$ satisfies the requirements of our lemma, in virtue of $A_n \supset A_{n-1} \supset \cdots \supset A_1 \supset A_0$ and $B_n \subset B_{n-1} \subset \cdots \subset B_1 \subset B_0$. In case, however, no such integer n exists, then after putting

$$\lim_{n=\infty} (A_n + \mathbf{T}_n) = A' + \mathbf{T}', \qquad A'\cdot\mathbf{T}' = 0 ,$$

and

$$\lim_{n=\infty} (\mathbf{T}_n + B_n) = \mathbf{T}' + B', \qquad \mathbf{T}' \cdot B' = 0,$$

it is easily seen that the sequence of normal representations $\{A_n + \mathbf{T}_n + B_n\}$ satisfies the hypothesis of Lemma 4. Consequently we get the normal representation

$$\mathbf{C} = A' + \mathbf{T}' + B'$$

which fulfills our requirements, since evidently $A' \supset A_0$ and $B' \subset B_0$.

LEMMA 6. *If G is an equicontinuous collection of simple closed curves, all of diameter $\geqq r > 0$, in a compact metric space W, then the closure $\bar{G}$ of G is a self-compact*[14] *equicontinuous collection of simple closed curves.*

Demonstration. Consider, in G, any infinite sequence of curves

(1) $$\mathbf{T}_1, \mathbf{T}_2, \cdots, \mathbf{T}_n, \cdots$$

that converges to a sequential limit L in W. Since, by hypothesis, W is compact and $\mathbf{T}_n$ is of diameter $\geqq r$, it follows that L is a compact continuum. We proceed to show that L is a simple closed curve.

Let a and b be an arbitrary pair of distinct points of L and choose on $\mathbf{T}_n$, for each value of n, points a_n and b_n so that the sequences $\{a_n\}$ and $\{b_n\}$ converge respectively to the points a and b. Denote by A_n and B_n the two arcs determined on $\mathbf{T}_n$ by the point pair (a_n, b_n). Then we can select a subsequence of curves

(2) $$\mathbf{T}_{n1}, \mathbf{T}_{n2}, \cdots, \mathbf{T}_{n_i}, \cdots$$

from the sequence (1) such that the corresponding arc sequences $\{A_{n_i}\}$ and $\{B_{n_i}\}$ converge respectively to sequential limits A and B in L. Clearly $A + B = L$ and $A \cdot B \supset a + b$.

Now suppose that $A \cdot B$ contains a third point c distinct from a and b. Then choose points P_i and Q_i upon A_{n_i} and B_{n_i} respectively in a fashion that the sequences $\{P_i\}$ and $\{Q_i\}$ both converge to the point c. Next, take ϵ, a positive number, less than both $\rho(a, c)$ and $\rho(b, c)$; then clearly all but a finite number of the arcs $P_i\, a_{n_i}\, Q_i$ and $P_i\, b_{n_i}\, Q_i$ on $\mathbf{T}_{n_i}$ are of diameter $> \epsilon$. Consequently we can choose curves $\mathbf{T}_{n_x}$, from (2), for which $\rho(P_x, Q_x)$ is arbitrarily small, while neither of the arcs joining P_x to Q_x in $\mathbf{T}_{n_x}$ is of diameter less than ϵ. This, however, implies the contradiction that the collection G is not equicontinuous; hence we conclude that $A \cdot B$ consists of but two points—a and b. Evidently then $L - (a + b)$ is not connected, and since the disconnecting point pair (a, b) was chosen *arbitrarily*, it follows that L is a simple closed curve.

It has been shown that the limit of any convergent sequence of curves of G is a simple closed curve in W. Hence $\bar{G} = G + \{L\}$ is a self-compact collection of curves and is obviously equicontinuous.

[14] A set G of curves in a compact metric space is said to be *self-compact* if the sequential limit of every convergent subsequence of G is also a curve of G.

II. Continua without Boundary Curves

THEOREM 1. *If a continuum of class $\mathfrak{R}$ does not have a boundary curve, then each of its proper cyclic elements is a simple closed surface.*

Demonstration. Assume that $\mathbf{C}$ is such a cyclic element; then according to Zippin,[15] it is sufficient to show that $\mathbf{C}$ satisfies non-vacuously the Jordan Curve Theorem: namely, that every simple closed curve in $\mathbf{C}$ determines in this set precisely two complementary domains of which it is the common boundary.

Let $\mathbf{T}_0$ be an arbitrary simple closed curve in $\mathbf{C}$.[16] Then since $\mathbf{C}$, by hypothesis, has no *boundary curves,* it follows that $\mathbf{C} - \mathbf{T}_0$ consists of more than a single component. Furthermore, there do not exist in $\mathbf{C} - \mathbf{T}_0$ three distinct components each of which has $\mathbf{T}_0$ as its boundary; for, the contrary implies (in contradiction to Lemma 3) that $\mathbf{C}$ does not have a normal representation with reference to $\mathbf{T}_0$. It now remains to be shown that the boundary of each component of $\mathbf{C} - \mathbf{T}_0$ is precisely $\mathbf{T}_0$ itself.

Suppose, on the contrary, that $\mathbf{C} - \mathbf{T}_0$ contains a component D having as its frontier $\mathbf{F}(D)$ a proper subset of $\mathbf{T}_0$. Then on taking a normal representation $\mathbf{C} = A_0 + \mathbf{T}_0 + B_0$, we may, without loss of generality, assume that D is a subset of B_0. Now, if P and Q are a pair of points bounding an interval on $\mathbf{T}_0$ that is complementary to the set $\mathbf{F}(D)$, it readily follows that $\mathbf{T}_0 + B_0$ is *separated* by $P + Q$; consequently, the demonstration of Theorem 1 is effected provided that we establish the following

FUNDAMENTAL LEMMA: *If $A_0 + \mathbf{T}_0 + B_0$ is a normal representation of $\mathbf{C}$ for which the point-set $\mathbf{T}_0 + B_0$ is separated by a point pair (P, Q) of $\mathbf{T}_0$, then $\mathbf{C}$ contains at least one boundary curve $\mathbf{T}$.*

Demonstration. Assume $A_0 + \mathbf{T}_0 + B_0$ to be a normal representation of $\mathbf{C}$ which satisfies our hypothesis. Then with $A_0 + \mathbf{T}_0 + B_0$ as an initial element, we proceed, in an inductive fashion, to obtain an infinite sequence $\mathbf{S}$ of normal representations of $\mathbf{C}$; this will aid us in establishing the existence of a boundary curve $\mathbf{T}$ in $\mathbf{C}$.

Take an infinite sequence of positive numbers $\epsilon_1, \epsilon_2, \cdots, \epsilon_n, \cdots$ such that

$$(1) \qquad \epsilon_n > 2\epsilon_{n+1}, \qquad n = 1, 2, 3, \cdots.$$

If then the n^{th} element $A_{n-1} + \mathbf{T}_{n-1} + B_{n-1}$ of $\mathbf{S}$ is given, satisfying the condition

(2) $\mathbf{T}_{n-1} + B_{n-1}$ is *separated* by the point-pair (P, Q) of $\mathbf{T}_{n-1}$, we determine the $(n + 1)^{\text{st}}$ element of $\mathbf{S}$ as follows:

The continuum $\mathbf{T}_{n-1} + B_{n-1}$, being peanian, can be covered by a finite collection $\mathfrak{S}^n$ of its regions U_i^n, where

$$(3) \qquad \delta(U_i^n) < \epsilon_n, \quad \text{for} \quad U_i^n \subset \mathfrak{S}^n.$$

[15] L. Zippin, On continuous curves and the Jordan Curve Theorem, Amer. Jour. Math., vol. 52 (1930), pp. 331–350. Cf. Theorem 3′, pp. 340–341.

[16] As $\mathfrak{C}$ is cyclic and distinct from a point, it evidently contains at least one simple closed curve.

Let us consider the collection $\mathbf{H}_n$ of normal representations $A_n^\sigma + \mathbf{T}_n^\sigma + B_n^\sigma$ of $\mathbf{C}$, each of which satisfies the conditions

a) $A_n^\sigma \supset A_{n-1}, \qquad B_n^\sigma \subset B_{n-1};$

b) $\mathbf{T}_n^\sigma + B_n^\sigma$ is *separated* by the point-pair (P, Q) of $\mathbf{T}$; and

c) B_n^σ does *not* contain any region U_i^n of $\mathfrak{S}^n$.

It is of utmost importance that we be convinced that the collection $\mathbf{H}_n$ is actually non-vacuous. In view of the finiteness of $\mathfrak{S}^n$, this can be shown in an elementary fashion with the aid of Lemma 5. We leave the proof to the reader and proceed to the following

DEFINITION. *If* $\mathbf{T}$ *is a simple closed curve and* ϵ_n *a positive number, then* $\xi_n(\mathbf{T})$ *denotes the greatest integer* r *such that* $\mathbf{T}$ *contains a collection of* r *non-overlapping*[17] *simple continuous arcs, each of diameter* $\geqq \epsilon_n$.

Now set

(4) $\qquad \pi_n = \text{lower bound } \xi_{n-1}(\mathbf{T}_n^\sigma), \text{ for } A_n^\sigma + \mathbf{T}_n^\sigma + B_n^\sigma \subset \mathbf{H}_n;$

and choose $A_n + \mathbf{T}_n + B_n$ an arbitrary normal representation of $\mathbf{H}_n$ for which $\xi_{n-1}(\mathbf{T}_n) = \pi_n$. We take the representation $A_n + \mathbf{T}_n + B_n$ as the $(n+1)^{\text{st}}$ element of the collection $\mathbf{S}$.

With the initial element $A_0 + \mathbf{T}_0 + B_0$ having been given, and in addition, a method for obtaining an $(n+1)^{\text{st}}$ normal representation when the n^{th} is determined, we generate an infinite sequence of normal representations $\mathbf{S} \equiv \{A_n + \mathbf{T}_n + B_n\}$ having the following properties:

A) $A_n \supset A_{n-1}, \quad B_n \subset B_{n-1}, \quad n \geqq 1;$

B) $\mathbf{T}_n + B_n$ is *separated* by the point-pair (P, Q) of $\mathbf{T}_n, \quad n \geqq 0;$

C) B_n does *not* contain any region U_i^n of $\mathfrak{S}^n, \quad n \geqq 1;$

D) $\xi_{n-1}(\mathbf{T}_n) = \pi_n, \quad n \geqq 1.$

Let $\{\mathbf{T}_n\}$ be the sequence of separating curves associated with the normal representations of $\mathbf{S}$. Then, for integers $l > m > n$, it follows, as a consequence of A), that

(5) $\qquad \mathbf{T}_m$ *separates* the point-sets $(\mathbf{T}_l - \mathbf{T}_m)$ and $(\mathbf{T}_n - \mathbf{T}_m)$ in $\mathbf{C}$.

We establish two properties of $\{\mathbf{T}_n\}$:

PROPERTY I: *The sequence* $\{\mathbf{T}_n\}$ *converges to a sequential limit.* For suppose $\{\mathbf{T}_{u_i}\}$ and $\{\mathbf{T}_{v_i}\}$ are subsequences of $\{\mathbf{T}_n\}$ that converge respectively to distinct sequential limits L and M in $\mathbf{C}$. Then one of the sets L and M, say M, contains a point T that does not belong to L. Take a region $U(T)$ in $\mathbf{C}$ of diameter $< \frac{1}{2}\rho(T, L)$ and let $\bar{n}$ be an integer such that for $v_i > \bar{n}$, $\mathbf{T}_{v_i}$ has points in $U(T)$.

[17] A collection of arcs L is said to be *non-overlapping* provided that no interior point of an arc of L belongs to any other arc of L.

Select also an integer $\bar{m}$ such that for $u_i > \bar{m}$ the upper distance of $\mathbf{T}_{u_i}$ from L is less than $\frac{1}{2}\rho(T, L)$; so that $\mathbf{T}_{u_i}$ has *no* point in $U(T)$. Now choose three integers u_a, u_b, u_c satisfying $u_a < \bar{m}, v_b > \bar{n} < v_c, v_b < u_a < v_c$.

It appears then that $\mathbf{T}_{u_a}$ does not separate $\mathbf{C}$ between $(\mathbf{T}_{v_b} - \mathbf{T}_{u_a})$ and $(\mathbf{T}_{v_c} - \mathbf{T}_{u_a})$ since both these sets have points in the region $U(T)$, while $\mathbf{T}_{u_a} \cdot U(T) = 0$. We have then a contradiction of 5); and Property I is established.

Property II. *The curves of* $\{\mathbf{T}_n\}$ *form an equicontinuous collection.* Suppose that this is not so; then it readily follows that there exists an infinite subsequence $\{\mathbf{T}_{n_i}\}$ of $\{\mathbf{T}_n\}$, and in $\mathbf{C}$, an infinite sequence of arcs $\{a_i b_i\}$ such that

1°) $n_i < n_{i+1}, \quad i = 1, 2, 3, \cdots;$

2°) $(a_i b_i) \cdot \mathbf{T}_{n_i} = a_i + b_i, \quad i = 1, 2, 3, \cdots;$

3°) $\lim_{i=\infty} a_i b_i =$ a point T in $\mathbf{C}$; and

4°) The two arcs joining the points a_i and b_i on $\mathbf{T}_{n_i}$ can be denoted by $E^1_{n_i}$ and $E^2_{n_i}$ in such a way that the arc sequences $\{E^1_{n_i}\}$ and $\{E^2_{n_i}\}$ converge respectively to sequential limiting continua E^1 and E^2, each of diameter > 0.

This situation suggests a consideration of the following possibilities:

Case 1. The sequences $\{\mathbf{T}_{n_i}\}$ and $\{a_i b_i\}$ were so chosen as to satisfy the condition

$$5°) \qquad \widehat{a_i b_i} \subset B_{n_i}, \qquad i = 1, 2, 3, \cdots.$$

Since 1), the continuum $\mathbf{C}$ is cyclic, and 2), the continua E^1 and E^2 are of diameter > 0, it follows that a region U of $\mathbf{C}$ can be taken about the point T such that

(6) $\mathbf{C} - \bar{U}$ is a *connected* point-set, and

(7) $\delta(U) < \delta(E^u)$, $u = 1, 2$. As a consequence of 4°) and 7), there is an integer v_1 such that

(8) $(\mathbf{C} - \bar{U}) \cdot E^u_{n_v} \neq 0, \quad u = 1, 2, \quad$ when $v > v_1$.

For each integer h, let W_h denote the subcollection of $\mathfrak{S}^{n_h}$ which consists of all of those regions $U^{n_h}_i$ for which

(9) $(a_h b_h) \cdot U^{n_h}_i \neq 0$.

Then, by 1), 3), and 3°) there exists an integer v_2 such that

(10) $\rho(U^{n_v}_i, \mathbf{C} - \bar{U}) > 3\epsilon_{n_v - 1}$, when $U^{n_v}_i \subset W_v$ and $v > v_2$.

Now let v assume a fixed value greater than both v_1 and v_2, so that the relations 8) and 10) hold simultaneously. It readily follows from 6) and 8) that

the point-set $\mathbf{C} - \overline{U}$ contains an arc e_1e_2 that joins a point e_1 of $E^1_{n_v}$ to a point e_2 of $E^2_{n_v}$, and which has no other point on the curve $\mathbf{T}_{n_v} \equiv E^1_{n_v} + E^2_{n_v}$.

In consequence of C), 5°), and 9) it is evident that every region of the collection W_v must contain at least one point of the curve $\mathbf{T}_{n_v}$. Hence, since $\mathbf{T}_{n_v}$ is the sum of the two arcs $e_1a_ve_2$ and $e_1b_ve_2$, it follows that if $U^{n_v}_i$ is any region of W_v, then either

(11) $e_1 a_v e_2 \cdot U^{n_v}_i \neq 0$, or else

(12) $e_1 b_v e_2 \cdot U^{n_v}_i \neq 0$.

Accordingly, we have the representation $W_v = X + Y$, where a region of W_v belongs to X provided it satisfies the relation 11), and belongs to Y provided it satisfies the relation 12).

Since every point of the arc $a_v b_v$ lies in some region of the collection W_v, it is evident that the point-set consisting of the sum of the points of all the regions of W_v is *connected*. Hence there is some region $U^{n_v}_x$ of X that contains a point of (or is possibly identical with) a region $U^{n_v}_y$ of Y. Consider the point-set

(13) $V = U^{n_v}_x + U^{n_v}_y$ which is *open* and *connected*, and consequently a region. From 1), 3), and 13) we get

(14) $$\delta(V) < 2\epsilon_{n_v} < \epsilon_{n_v-1}.$$

There evidently exists in V an arc ab that joins a point a on arc $e_1 a_v e_2$ to a point b on arc $e_1 b_v e_2$, without having any other point on the curve $\mathbf{T}_{n_v}$. Then

(15) $\overset{\frown}{ab} \subset B_{n_v}$; for, since the point-pairs (a_v, b_v) and (e_1, e_2) separate each other on the *boundary curve* $\mathbf{T}_{n_v}$ of $(\mathbf{T}_{n_v} + B_{n_v})$, it follows from relation 5^0) and Lemma 1 that $\overset{\frown}{e_1e_2} \subset A_{n_v}$; consequently, since also the point-pairs (a, b) and (e_1, e_2) separate each other on $\mathbf{T}_{n_v}$, we conclude in an analogous fashion that $\overset{\frown}{ab} \subset B_{n_v}$.

Let us consider the two arcs ae_1b and ae_2b upon $\mathbf{T}_{n_v}$. Since, by B), the point-set $\mathbf{T}_{n_v} + B_{n_v}$ is separated by the point-pair (P, Q) of $\mathbf{T}_{n_v}$, it may readily be shown, with the aid of 15), that one of the arcs (ae_1b, ae_2b), say ae_1b, is such that if

(16) $\mathbf{T}^\sigma_{n_v} = ab + ae_1b$, then P and Q are points of $\mathbf{T}^\sigma_{n_v}$, and

(17) $\mathbf{T}^\sigma_{n_v} - (P + Q)$ is *not* a subset of a *single* component of $\mathbf{T}_{n_v} + B_{n_v} - (P + Q)$.

Now, since $\mathbf{T}^\sigma_{n_v} \subset \mathbf{T}_{n_v} + B_{n_v}$, there exists, according to Lemma 5, a normal representation

(18) $\mathbf{C} = A^\sigma_{n_v} + \mathbf{T}^\sigma_{n_v} + B^\sigma_{n_v}$, where $A^\sigma_{n_v} \supset A_{n_v}$, and $B^\sigma_{n_v} \subset B_{n_v}$.

The relation A), with n replaced by n_v, yields $A_{n_v} \supset A_{n_v-1}$ and $B_{n_v} \supset B_{n_v-1}$, so that, by 18) we get

a′) $A^{\sigma}_{n_v} \supset A_{n_v-1}$, $B^{\sigma}_{n_v} \subset B_{n_v-1}$. Also, from
$\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \subset \mathbf{T}_{n_v} + B_{n_v}$ and 17) we observe that $\mathbf{T}^{\sigma}_{n_v} - (P + Q)$ is *not* a subset of a *single* component of $\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} - (P + Q)$, which implies that

b′) $\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v}$ is *separated* by the point-pair (P, Q) of $\mathbf{T}^{\sigma}_{n_v}$.

Furthermore, the relation c), together with $B^{\sigma}_{n_v} \subset B_{n_v}$, insures that

c′) $B^{\sigma}_{n_v}$ does *not* contain any region $U^{n_v}_i$ of $\mathfrak{S}^{n_v}$.

In view of a′), b′), and c′) we conclude that the normal representation

$$A^{\sigma}_{n_v} + \mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \text{ is a member of the collection } \mathbf{H}_{n_v}\,.$$

Now let L be a collection consisting of $\xi_{n_v-1}\,(\mathbf{T}^{\sigma}_{n_v})$ non-overlapping arcs, of diameter $\geqq \epsilon_{n_v-1}$, lying in the curve $\mathbf{T}^{\sigma}_{n_v}$. From 14) and $ab \subset V$ we get $\delta(ab) < \epsilon_{n_v-1}$.
Consequently, it appears that at *most* 2 of the arcs of L have points upon the arc ab, implying, by 16), that at *least* $\xi_{n_v-1}\,(\mathbf{T}^{\sigma}_{n_v}) - 2$ of the arcs of L lie within arc ae_1b. Since, however, a and b are points of the region V, while e_2 is a point of $\mathbf{C} - \overline{U}$, it follows from 10) and 13) that ae_2b is an arc of diameter $> 3\,\epsilon_{n_v-1}$, implying that ae_2b contains a set consisting of at *least* 3 non-overlapping simply continuous arcs. Then, from $\mathbf{T}_{n_v} = ae_1b + ae_2b$, we see finally that $\xi_{n_v-1}\,(\mathbf{T}_{n_v}) \geqq [\xi_{n_v-1}\,(\mathbf{T}^{\sigma}_{n_v}) - 2] + 3 = \xi_{n_v-1}\,(\mathbf{T}^{\sigma}_{n_v}) + 1$, which, in view of the relations 4) and $A^{\sigma}_{n_v} + \mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \subset \mathbf{H}_{n_v}$, is a contradiction of $\xi_{n-1}\,(\mathbf{T}_n) = \pi_n$, for $n = n_v$. This contradiction disposes of Case 1, and we turn to

Case 2: It is *impossible* to choose the sequences $\{\mathbf{T}_{n_i}\}$ and $\{a_ib_i\}$ so as to satisfy the condition 5°) of Case 1.

Then we may assume, without loss of generality, that $\{\mathbf{T}_{n_i}\}$ and $\{a_ib_i\}$ were so chosen that

$$(19) \qquad \widehat{a_ib_i} \subset A_{n_i}\,, \qquad i = 1, 2, 3, \cdots .$$

As a further consequence of our hypothesis it can be shown that the two continua E^1 and E^2 have but *one* common point, namely the point T. Take T_μ an arbitrary point of $E^\mu - T$, $u = 1, 2$; and about T_μ let U_μ be a region of $\mathbf{C}$ such that

$$(20) \qquad \overline{U}_1 \cdot E^2 = \overline{U}_2 \cdot E^1 = 0\,.$$

In view of 3°), 4°), and 20), there exists an integer k such that for $r > k$ we have with reference to the simple closed curve $\mathbf{T}_{n_r}, = E^1_{n_r} + E^2_{n_r}$, the relations

$$(21) \qquad \mathbf{T}_{n_r} \cdot U_\mu = E^\mu_{n_r} \cdot U_\mu \neq 0\,, \qquad \mu = 1, 2, \text{ and}$$

(22) $a_r b_r \subset \mathbf{C} - (\bar{U}_1 + \bar{U}_2)$. Let e_1 and e_2 be points of the curve $\mathbf{T}_{nk}$ that ie, respectively, in the regions U_1 and U_2. Then

$$\mathbf{T}_{nk} = e_1 a_k e_2 + e_1 b_k e_2$$

and for each integer $x > k$ we have

(23) $$e_1 a_k e_2 \cdot a_x b_x \neq 0 \neq e_1 b_k e_2 \cdot a_x b_x, \text{ where } a_x b_x \subset \{a_i b_i\};$$

for, if either of the arcs $e_1 a_k e_2$ and $e_1 b_k e_2$ should *fail* to contain a point of arc $a_x b_x$ when $x > k$, then in virtue of $\mathbf{T}_{nk} \subset A_{n_x} + \mathbf{T}_{n_x}$, 20)), 21)), and 22)), it would follow that an arc $c_x d_x$ can be determined in $A_{n_x} + \mathbf{T}_{n_x}$, having no point on arc $a_x b_x$ and only its end points on $\mathbf{T}_{n_x}$, in a fashion that the point pairs (a_x, b_x) and (c_x, d_x) separate each other on $\mathbf{T}_{n_x}$. Then, by 19) and condition (a) of Lemma 1, we arrive at the contradiction that $\mathbf{T}_{n_x}$ is not a *boundary curve* of the continuum $A_{n_x} + \mathbf{T}_{n_x}$.

As a consequence of the validity of 23) for all integers $x > k$ we observe, from $\lim_{i=\infty} a_i b_i = T$, (see 3°), that T is a point of each of the arcs $e_1 a_k e_2$ and $e_1 b_k e_2$. However, the relations $U_1 \supset e_1$, $U_2 \supset e_2$, $E^1 \supset T \subset E^2$, and 20) imply that the point T is an *interior* point of each of these arcs. We are, therefore, confronted with the contradiction that the simple closed curve $\mathbf{T}_{nk} = e_1 a_k e_2 + e_1 b_k e_2$ has a *singular* point, namely, the point T.

The contradictions that have been obtained in the treatment of Cases 1 and 2 establish, without doubt, the *equicontinuity* of the collection $\{\mathbf{T}_n\}$.

In virtue of Properties I and II, it follows, by Lemma 6, that the sequential limit of $\{\mathbf{T}_n\}$ is a simple closed curve $\mathbf{T}$ of $\mathbf{C}$. Also, the relation A) insures that the sequences $\{A_n + \mathbf{T}_n\}$ and $\{\mathbf{T}_n + B_n\}$ both converge to *sequential limits* which we denote respectively by $A + \mathbf{T}$ and $\mathbf{T} + B$, under the assumption that $A \cdot \mathbf{T} = \mathbf{T} \cdot B = 0$. By Lemma 4, it follows then that $A + \mathbf{T} + B$ is a normal representation of $\mathbf{C}$. Furthermore, the point-set B is *vacuous*, for it can be easily shown that $B \neq 0$ implies a contradiction of the condition C) for a *sufficiently large* value of n. Consequently, we have the normal representation $\mathbf{C} = A + \mathbf{T}$, which establishes $\mathbf{T}$ as a *boundary curve* of $\mathbf{C}$. This completes the proof of our Fundamental Lemma.

As a direct consequence of Theorem 1 we get the following

COROLLARY. *A simple closed surface may be characterized as a cyclic continuum of class $\mathfrak{R}$, distinct from a point, which does not have a boundary curve.*

III. The Boundary Curves of $\mathbf{M}_1$

We prove two theorems concerning these curves:

THEOREM 2. *The collection G^r, consisting of the boundary curves in* $\mathbf{M}$ *of diameter $\geqq r$, is a self-compact equicontinuous collection.*

Demonstration. We show first that G^r is equicontinuous. Suppose that this is not so; then it follows that G^r contains an infinite sequence of curves

(1) $$\mathbf{T}_1, \mathbf{T}_2, \cdots, \mathbf{T}_n, \cdots$$

The relation A), with n replaced by n_v, yields $A_{n_v} \supset A_{n_v-1}$ and $B_{n_v} \supset B_{n_v-1}$, so that, by 18) we get

a′) $A^{\sigma}_{n_v} \supset A_{n_v-1}$, $B^{\sigma}_{n_v} \subset B_{n_v-1}$. Also, from $\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \subset \mathbf{T}_{n_v} + B_{n_v}$ and 17) we observe that $\mathbf{T}^{\sigma}_{n_v} - (P + Q)$ is *not* a subset of a *single* component of $\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} - (P + Q)$, which implies that

b′) $\mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v}$ is *separated* by the point-pair (P, Q) of $\mathbf{T}^{\sigma}_{n_v}$.

Furthermore, the relation c), together with $B^{\sigma}_{n_v} \subset B_{n_v}$, insures that

c′) $B^{\sigma}_{n_v}$ does *not* contain any region $U^{n_v}_i$ of $\mathfrak{S}^{n_v}$.

In view of a′), b′), and c′) we conclude that the normal representation

$$A^{\sigma}_{n_v} + \mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \text{ is a member of the collection } \mathbf{H}_{n_v}.$$

Now let L be a collection consisting of $\xi_{n_v-1}(\mathbf{T}^{\sigma}_{n_v})$ non-overlapping arcs, of diameter $\geqq \epsilon_{n_v-1}$, lying in the curve $\mathbf{T}^{\sigma}_{n_v}$. From 14) and $ab \subset V$ we get $\delta(ab) < \epsilon_{n_v-1}$.
Consequently, it appears that at *most* 2 of the arcs of L have points upon the arc ab, implying, by 16), that at *least* $\xi_{n_v-1}(\mathbf{T}^{\sigma}_{n_v}) - 2$ of the arcs of L lie within arc ae_1b. Since, however, a and b are points of the region V, while e_2 is a point of $\mathbf{C} - \overline{U}$, it follows from 10) and 13) that ae_2b is an arc of diameter $> 3\,\epsilon_{n_v-1}$, implying that ae_2b contains a set consisting of at *least* 3 non-overlapping simply continuous arcs. Then, from $\mathbf{T}_{n_v} = ae_1b + ae_2b$, we see finally that $\xi_{n_v-1}(\mathbf{T}_{n_v}) \geqq [\xi_{n_v-1}(\mathbf{T}^{\sigma}_{n_v}) - 2] + 3 = \xi_{n_v-1}(\mathbf{T}^{\sigma}_{n_v}) + 1$, which, in view of the relations 4) and $A^{\sigma}_{n_v} + \mathbf{T}^{\sigma}_{n_v} + B^{\sigma}_{n_v} \subset \mathbf{H}_{n_v}$, is a contradiction of $\xi_{n-1}(\mathbf{T}_n) = \pi_n$, for $n = n_v$. This contradiction disposes of Case 1, and we turn to

Case 2: It is *impossible* to choose the sequences $\{\mathbf{T}_{n_i}\}$ and $\{a_ib_i\}$ so as to satisfy the condition 5°) of Case 1.

Then we may assume, without loss of generality, that $\{\mathbf{T}_{n_i}\}$ and $\{a_ib_i\}$ were so chosen that

$$\widehat{a_ib_i} \subset A_{n_i}, \qquad i = 1, 2, 3, \cdots. \tag{19}$$

As a further consequence of our hypothesis it can be shown that the two continua E^1 and E^2 have but *one* common point, namely the point T. Take T_μ an arbitrary point of $E^\mu - T$, $u = 1, 2$; and about T_μ let U_μ be a region of $\mathbf{C}$ such that

$$\overline{U}_1 \cdot E^2 = \overline{U}_2 \cdot E^1 = 0. \tag{20}$$

In view of 3°), 4°), and 20), there exists an integer k such that for $r > k$ we have with reference to the simple closed curve $\mathbf{T}_{n_r}$, $= E^1_{n_r} + E^2_{n_r}$, the relations

$$\mathbf{T}_{n_r} \cdot U_\mu = E^\mu_{n_r} \cdot U_\mu \neq 0, \qquad \mu = 1, 2, \text{ and} \tag{21}$$

(22) $a_r b_r \subset \mathbf{C} - (\overline{U}_1 + \overline{U}_2)$. Let e_1 and e_2 be points of the curve $\mathbf{T}_{nk}$ that ie, respectively, in the regions U_1 and U_2. Then

$$\mathbf{T}_{nk} = e_1 a_k e_2 + e_1 b_k e_2$$

and for each integer $x > k$ we have

(23) $$e_1 a_k e_2 \cdot a_x b_x \neq 0 \neq e_1 b_k e_2 \cdot a_x b_x, \text{ where } a_x b_x \subset \{a_i b_i\};$$

for, if either of the arcs $e_1 a_k e_2$ and $e_1 b_k e_2$ should *fail* to contain a point of arc $a_x b_x$ when $x > k$, then in virtue of $\mathbf{T}_{nk} \subset A_{n_x} + \mathbf{T}_{n_x}$, 20), 21), and 22), it would follow that an arc $c_x d_x$ can be determined in $A_{n_x} + \mathbf{T}_{n_x}$, having no point on arc $a_x b_x$ and only its end points on $\mathbf{T}_{n_x}$, in a fashion that the point pairs (a_x, b_x) and (c_x, d_x) separate each other on $\mathbf{T}_{n_x}$. Then, by 19) and condition (a) of Lemma 1, we arrive at the contradiction that $\mathbf{T}_{n_x}$ is not a *boundary curve* of the continuum $A_{n_x} + \mathbf{T}_{n_x}$.

As a consequence of the validity of 23) for all integers $x > k$ we observe, from $\lim_{i=\infty} a_i b_i = T$, (see 3°), that T is a point of each of the arcs $e_1 a_k e_2$ and $e_1 b_k e_2$. However, the relations $U_1 \supset e_1$, $U_2 \supset e_2$, $E^1 \supset T \subset E^2$, and 20) imply that the point T is an *interior* point of each of these arcs. We are, therefore, confronted with the contradiction that the simple closed curve $\mathbf{T}_{nk} = e_1 a_k e_2 + e_1 b_k e_2$ has a *singular* point, namely, the point T.

The contradictions that have been obtained in the treatment of Cases 1 and 2 establish, without doubt, the *equicontinuity* of the collection $\{\mathbf{T}_n\}$.

In virtue of Properties I and II, it follows, by Lemma 6, that the sequential limit of $\{\mathbf{T}_n\}$ is a simple closed curve $\mathbf{T}$ of $\mathbf{C}$. Also, the relation A) insures that the sequences $\{A_n + \mathbf{T}_n\}$ and $\{\mathbf{T}_n + B_n\}$ both converge to *sequential limits* which we denote respectively by $A + \mathbf{T}$ and $\mathbf{T} + B$, under the assumption that $A \cdot \mathbf{T} = \mathbf{T} \cdot B = 0$. By Lemma 4, it follows then that $A + \mathbf{T} + B$ is a normal representation of $\mathbf{C}$. Furthermore, the point-set B is *vacuous*, for it can be easily shown that $B \neq 0$ implies a contradiction of the condition C) for a *sufficiently large* value of n. Consequently, we have the normal representation $\mathbf{C} = A + \mathbf{T}$, which establishes $\mathbf{T}$ as a *boundary curve* of $\mathbf{C}$. This completes the proof of our Fundamental Lemma.

As a direct consequence of Theorem 1 we get the following

COROLLARY. *A simple closed surface may be characterized as a cyclic continuum of class* $\mathfrak{R}$, *distinct from a point, which does not have a boundary curve.*

III. The Boundary Curves of $\mathbf{M}_1$

We prove two theorems concerning these curves:

THEOREM 2. *The collection* G^r, *consisting of the boundary curves in* $\mathbf{M}$ *of diameter* $\geqq r$, *is a self-compact equicontinuous collection.*

Demonstration. We show first that G^r is equicontinuous. Suppose that this is not so; then it follows that G^r contains an infinite sequence of curves

(1) $$\mathbf{T}_1, \mathbf{T}_2, \cdots, \mathbf{T}_n, \cdots$$

that contain, respectively, point pairs

$$(P_1, Q_1), (P_2, Q_2), \cdots, (P_n, Q_n), \cdots$$

such that

(i) $\lim_{n=\infty} \rho(P_n, Q_n) = 0$, and

(ii) both arcs joining P_n to Q_n on $\mathbf{T}_n$ are of diameter $> \epsilon > 0$.

Let A_n and B_n denote the two arcs determined on $\mathbf{T}_n$ by the point pair (P_n, Q_n); then we can choose from the sequence (1) a subsequence $\{\mathbf{T}_{n_i}\}$ such that the corresponding sequences $\{A_{ni}\}$, $\{B_{ni}\}$, and $\{P_{ni}\}$ converge respectively to sequential limits A, B, and T. The relation (i) now implies that the sequence $\{Q_{n_i}\}$ also converges to T. The relation (ii) implies, furthermore, that the continua A and B are both of diameter > 0, and hence contain, respectively, points a and b each distinct from the point T. Now select, for each value of i, a point a_i on A_{n_i} and a point b_i on B_{n_i}, so that the sequence $\{a_i\}$ converges to the point a, while the sequence $\{b_i\}$ converges to the point b.

Clearly all but a finite number of the curves of $\{\mathbf{T}_{n_i}\}$ are a subset of a single cyclic element C of $\mathbf{M}$; so that the points a, b, and T belong to C. Take in C a region R containing T for which $\mathsf{C} - \bar{R} \supset a + b$. Then, as T is not a cut-point of C, there exists about T a sub-region U of R such that $\mathsf{C} - U$ is a *peanian* continuum.[18] Since all but a finite number of the point pairs (a_i, b_i) lie in $\mathsf{C} - U$, and all but a finite number of the point pairs (P_{n_i}, Q_{n_i}) lie in U, it follows that for a sufficiently large integer v, we get

$$a_v + b_v \subset \mathsf{C} - U \quad \text{and} \quad P_{n_v} + Q_{n_v} \subset U .$$

In the peanian continuum $\mathsf{C} - U$ take an arc $a_v b_v$ joining a_v to b_v; and in the region U, take an arc $P_{n_v} Q_{n_v}$ joining P_{n_v} to Q_{n_v}. The arc $a_v b_v$ clearly contains a sub-arc $\alpha\beta$ that joins a point α of A_{n_v} to a point β of B_{n_v} and which has no other points on $\mathbf{T}_{n_v}$.

Now consider the two arcs $\alpha P_{n_v}\beta$ and $\alpha Q_{n_v}\beta$ that are determined on $\mathbf{T}_{n_v}$ by the point pair (α, β). The arc $P_{n_v} Q_{n_v}$ contains a sub-arc $\gamma\delta$ that joins a point γ on $\alpha P_{n_v}\beta$ to a point δ on $\alpha Q_{n_v}\beta$ and which has no other points on $\mathbf{T}_{n_v}$.

The two arcs $\alpha\beta$ and $\gamma\delta$ evidently have no common points, and furthermore, the point pairs (α, β) and (γ, δ) separate each other on $\mathbf{T}_{n_v}$. Consequently it follows from condition (a) of Lemma 1 that $\mathbf{T}_{n_v}$ is not a *boundary curve* of $\mathbf{M}$; and in view of this contradiction we conclude that the collection G^r must be equicontinuous.

It is now shown that G^r is a self-compact collection of curves. Consider an arbitrary sequence $\{\mathbf{T}_n\}$ of curves of G^r that converges to a sequential limiting

[18] Cf. C. Kuratowski, Une caractérisation topologique de la surface de la sphère, Fund. Math., vol. 13 (1929), pp. 307–318; see *property* (γ), pp. 314–315; and note 2), p. 315.

set in $\mathbf{M}$; then it follows, from Lemma 6, that this limiting set is a simple closed curve $\mathbf{T}$. We proceed to identify $\mathbf{T}$ as a boundary curve of $\mathbf{M}$.

Let $\mathbf{C}$ be the cyclic element of $\mathbf{M}$ that contains $\mathbf{T}$; then there is evidently an integer μ such that $\mathbf{T}_n \subset \mathbf{C}$ for $n > \mu$. Furthermore, since $\mathbf{T}_n$, $n > \mu$, is, by hypothesis, a boundary curve of $\mathbf{M}$, and hence of $\mathbf{C}$, we clearly have the normal representation $\mathbf{C} = A_n + \mathbf{T}_n$. Now consider the infinite sequence of normal representations $\{A_n + \mathbf{T}_n\}_{n>\mu}$ of $\mathbf{C}$. This sequence evidently satisfies the hypothesis of Lemma 4; consequently, on setting

$$\lim_{n=\infty} A_n + \mathbf{T}_n = A + \mathbf{T},$$

where $A \cdot \mathbf{T} = 0$, we get the normal representation $\mathbf{C} = A + \mathbf{T}$. This implies that $\mathbf{T}$ is a *boundary curve* of $\mathbf{C}$, and hence of $\mathbf{M}$ (since $\mathbf{C}$ is a cyclic element of $\mathbf{M}$). It now follows, by definition, that G^r is self-compact.

THEOREM 3. *If* $\{\mathbf{T}_n\}$ *is a sequence of boundary curves of* $\mathbf{M}$ *that converges to a boundary curve* $\mathbf{T}$ *as a sequential limit, then the set of points, each of which is common to all but a finite number of the curves of* $\{\mathbf{T}_n\}$, *is dense on* $\mathbf{T}$.

Demonstration. We see from Theorem 2 that the sequence $\{\mathbf{T}_n\}$ is equicontinuous. This, with the aid of our hypothesis that $\mathbf{T}$ and $\mathbf{T}_n$, $n = 1, 2, \cdots$, are *boundary curves,* easily leads to the conclusion that if D is any component of $\mathbf{C} - \mathbf{T}$, then there exists an integer n_D such that

$$\text{(1)} \qquad \mathbf{F}(D) \subset \mathbf{T}_n, \quad \text{for} \quad n > n_D .$$

If, furthermore, B is any segment on $\mathbf{T}$ that contains no point of $\mathbf{F}(\mathbf{C} - \mathbf{T})$ then it is clear that there exists an integer n_B such that

$$B \subset \mathbf{T}_n, \qquad \text{for} \qquad n > n_B .$$

Now let L be an arbitrary segment on $\mathbf{T}$. If for some component D of $\mathbf{C} - \mathbf{T}$ we have $\mathbf{F}(D) \cdot L \neq 0$, then it follows, from (1), that the segment L contains at least one point that is common to all but a finite number of the curves $\{\mathbf{T}_n\}$. If, however, for each component D of $\mathbf{C} - \mathbf{T}$, we have

$$\text{(2)} \qquad \mathbf{F}(D) \cdot L = 0,$$

then, since L is an *open* subset of $\mathbf{T}$, it follows from Lemma 2 that relation (2) implies $\mathbf{F}(\mathbf{C} - \mathbf{T}) \cdot L = 0$. We conclude then that L is actually contained in all but a finite number of the curves $\{\mathbf{T}_n\}$. It has now been shown that an *arbitrary* segment L on $\mathbf{T}$ contains points common to all but a finite number of the curves $\{\mathbf{T}_n\}$; hence the set of such points is *dense* on $\mathbf{T}$.

IV. The Extension of the Property of Belonging to Class $\mathfrak{N}$

THEOREM E_1. *Let* $C, E_1, E_2, \cdots, E_n, \cdots$, *be an infinite sequence of peanian continua such that*

(a) $C_n = C + \sum_{i=1}^{n} E_i$, *is a continuum of class* $\mathfrak{N}$, $\} \; n = 1, 2, \cdots,$

(b) $C \cdot E_n \neq 0$,

(c) $\lim_{n=\infty} \delta(E_n) = 0$;

then $C_\omega = C + \sum_{i=1}^{\infty} E_i$ *is a continuum of class* $\mathfrak{R}$.

Demonstration. Let ϵ be an arbitrary positive number. Then, according to condition (c) of our hypothesis, there exists an integer r such that

$$(1) \qquad \delta(E_\mu) < \epsilon/3, \qquad \text{for} \qquad \mu > r.$$

Since the continuum C_r is peanian, it can be represented as the sum of a finite number of its peanian subcontinua

$$L_1^r, L_2^r, \cdots, L_s^r,$$

where

$$(2) \qquad \delta(L_i^r) < \epsilon/3, \qquad i = 1, 2, \cdots, s.$$

Let us consider the continuum C_μ for $\mu > r$. Evidently

$$(3) \qquad C_\mu = L_1^r + L_2^r + \cdots + L_s^r + E_{r+1} + \cdots + E_\mu.$$

Now let $H_i^{r,\mu}$, $(1 \leqq i \leqq s)$, denote the sum of those continua among $E_{r+1}, \cdots, E_\mu$ each of which has a point in common with L_i^r. Then, on setting

$$(4) \qquad L_i^\mu = L_i^r + H_i^{r,\mu}, \qquad i = 1, 2, \cdots, s,$$

we see from (3) and (4) that

$$(5) \qquad C_\mu = L_1^\mu + L_2^\mu + \cdots + L_s^\mu, \qquad \mu > r,$$

where the continua L_i^μ are evidently *peanian*; and

$$(6) \qquad \delta(L_i^\mu) < \epsilon, \qquad i = 1, 2, \cdots, s,$$

because of (1), (2), and (4).

Let us now put

$$L_i^\omega = \sum_{\mu=r+1}^{\infty} L_i^\mu, \qquad i = 1, 2, \cdots, s;$$

then evidently

$$(7) \qquad C_\omega = L_1^\omega + L_2^\omega + \cdots + L_s^\omega,$$

where again by (1), (2), and (4) we have

$$(8) \qquad \delta(L_i^\omega) < \epsilon, \qquad i = 1, 2, \cdots, s.$$

The relations (7) and (8) imply that the obviously compact point-set C_ω is in fact a *peanian* continuum. By a precisely similar argument it follows that L_i^ω, $(1 \leqq i \leqq s)$, is also a peanian continuum.

We now show that the continuum C_ω does not contain a *primitive skew curve.*

Suppose, on the contrary, that C_ω does contain such a curve Π having vertices $P_1, P_2, P_3, \cdots$ and edges $P_1P_2, P_1P_3, P_1P_4, \cdots$. Then about the respective vertices $P_1, P_2, P_3, \cdots$ choose regions $U_1, U_2, U_3, \cdots$ in C_ω such that

i) $\rho(U_x, U_y) > 0$ for each pair of distinct regions, and

ii) $\bar{U}_x$ contains no point on any edge of Π not on the vertex P_x (for each value of x).

We now assume that the positive number ϵ (which was introduced at the beginning of this demonstration) is subject to the following restrictions

(9) $\epsilon < \frac{1}{2}\rho(U_x, U_y)$ for each pair of distinct regions,

(10) the region U_x contains all points in C_ω at a distance $< \epsilon$ from P_x (for each value of x),

(11) if p and q are points of $\Pi - \sum U_x$ that lie respectively on *distinct* edges of Π, then $\epsilon < \frac{1}{2}\rho(p, q)$.

For each edge P_xP_y of Π, let $T^\omega_{x,y}$ denote the point-set consisting of those continua L^ω_i, appearing in (7), that have a point in common with P_xP_y. The set $T^\omega_{x,y}$ is evidently a peanian continuum containing the two vertices P_x and P_y. If the edges P_iP_j and P_kP_l do not have a common vertex, then, from (8), (9), and (11), we get

$$(12) \qquad T^\omega_{i,j}\cdot T^\omega_{k,l} = 0,$$

while if P_iP_j and P_jP_k have just one common vertex, then, from (8), (9), (10), and (11), we get

$$(13) \qquad P_j \subset T^\omega_{i,j}\cdot T^\omega_{j,k} \subset U_j.$$

Now choose an integer $m \geqq r$ in a fashion that

(14) every vertex P_x of Π is contained in the continuum C_m;

and for each integer $\nu \geqq m$ let $T^\nu_{x,y}$ denote the point-set consisting of those continua L^ν_i appearing in (5), when μ is replaced by ν, that have a point in common with the edge P_xP_y of Π. In general the closed point-set $T^\nu_{x,y}$ *will not be connected* for all values of $\nu \geqq m$; however, with the aid of condition (b) of our hypothesis, it follows that an integer $m_{x,y}$ exists such that

$$(15) \qquad T^\nu_{x,y} \text{ is } \textit{connected} \text{ for } \nu \geqq m_{x,y};$$

(to assume that $m_{x,y}$ does not exist will lead to the contradiction that $T^\omega_{x,y}$ is not connected). Now let $\bar{m}$ be the greatest of the integers $m_{x,y}$ associated with the edges P_xP_y of Π. Then each of the point-sets $T^{\bar{m}}_{x,y}$ is *connected*, and, moreover, is a *peanian* continuum (since each continuum $L^{\bar{m}}_i$ is peanian).

Since $T^{\bar{m}}_{x,y} \subset T^\omega_{x,y}$, the relation (12) implies

$$(16) \qquad T^{\bar{m}}_{i,j}\cdot T^{\bar{m}}_{k,l} = 0, \qquad \text{when} \qquad P_iP_j\cdot P_kP_l = 0;$$

while, from (13) and (14), we get

$$(17)\qquad P_i \subset T^{\bar m}_{i,j}\cdot T^{\bar m}_{j,k} \subset U_j, \qquad \text{when} \qquad P_iP_j\cdot P_jP_k = P_j.$$

In each continuum $T^{\bar m}_{x,y}$ let us take, finally, an arc $t^{\bar m}_{x,y}$ joining P_x to P_y. Then consider the point-set $\Pi_{\bar m} = \sum t^{\bar m}_{x,y}$. With the aid of (16) and (17), it readily follows that $\Pi_{\bar m}$ contains a primitive skew curve Ω.[19] Furthermore, Ω is evidently a subset of $C_{\bar m}$. This conclusion is clearly a contradiction of condition (a) of our hypothesis; therefore our theorem is proved.

Theorem E_2. *If* $\mathbf{M}_1$ *and* $\mathbf{M}_2$ *are continua of class* $\mathfrak{R}$ *such that* $\mathbf{M}_1\cdot\mathbf{M}_2$ *is a boundary curve of both* $\mathbf{M}_1$ *and* $\mathbf{M}_2$*, then* $\mathbf{M}_1 + \mathbf{M}_2$ *is a continuum of class* $\mathfrak{R}$.

Demonstration. We assume our proposition false, implying then that the peanian continuum $\mathbf{M}_1 + \mathbf{M}_2$ contains a *primitive skew curve* Π. Let $\mathbf{T}$ denote the simple closed curve $\mathbf{M}_1\cdot\mathbf{M}_2$, and consider the point-set $\Pi - \mathbf{T}$:

Case 1: $\Pi - \mathbf{T}$ consists of but a *finite* number of components. It readily follows then that the continuum $\Pi + \mathbf{T}$ contains at *most* a finite number of simple closed curves, and is in fact a *linear graph.*

Putting

$$\Gamma_\mu = \Pi\cdot\mathbf{M}_\mu + \mathbf{T}, \qquad \mu = 1, 2,$$

we observe that Γ_1 and Γ_2 are each linear graphs having $\mathbf{T}$ as a *boundary curve* and *failing* to contain a *primitive skew curve.* Consequently, by Theorem A, there exists, for $\mu = 1, 2$, a topological transformation $t_\mu(P)$ carrying Γ_μ into a subset $\Gamma'_\mu = t_\mu(\Gamma_\mu)$ of a spherical surface S. The image $t_\mu(\mathbf{T})$ of $\mathbf{T}$ in Γ'_μ is denoted by $\mathbf{T}'_\mu$, $\mu = 1, 2$.

At this point we introduce a plausible proposition on linear graphs, the proof of which we omit.

Proposition K. *If* H *is a linear graph lying in a spherical surface* S*, and if* I *is a boundary curve of* H*, then there exists in* S *a topological image* $\chi(H)$ *of* H *such that the simple closed curve* $\chi(I)$ *constitutes the boundary of one of the components of* $S - \chi(H)$.

In view of Proposition K, let $\chi_\mu(y)$, $\mu = 1, 2$, be a topological transformation of Γ'_μ into a subset $\Gamma''_\mu = \chi_\mu(\Gamma'_\mu)$ of S such that the simple closed curve $\mathbf{T}''_\mu = \chi_\mu(\mathbf{T}'_\mu)$ constitutes the *boundary* of a component I_μ of $S - \Gamma''_\mu$. On setting $h_\mu(P) = \chi_\mu[t_\mu(P)]$, $\mu = 1, 2$, it is clear that the transformation $h(Q) = h_1[h_2^{-1}(Q)]$ is, for $Q \subset \mathbf{T}''_2$, a homeomorphism between the two simple closed curves $\mathbf{T}''_1$ and $\mathbf{T}''_2$. It is furthermore a well known property of a spherical surface that $h(Q)$ can be *extended* to a homeomorphism $H(Q)$ that carries the 2-cell $S - I_2$ into the 2-cell $I_1 + \mathbf{T}''_1$. Now consider the transformation

$$T(P)\begin{cases} = h_1(P), & \text{when} \quad P \subset \Gamma_1, \\ = H[h_2(P)], & \text{when} \quad P \subset \Gamma_2. \end{cases}$$

[19] It may be observed that $\Pi_{\bar m}$ does not necessarily contain a subset homeomorphic with Π unless Π is assumed in the beginning to be homeomorphic with the complex $\mathfrak{C}$ instead of the complex $\mathfrak{L}$. Cf. Introduction.

It is readily seen that $T(P)$ is a topological transformation carrying $\Gamma_1 + \Gamma_2$ into a subset $T(\Gamma_1 + \Gamma_2)$ of the spherical surface S. But the *primitive skew curve* Π is a subset of $\Gamma_1 + \Gamma_2$; consequently we are lead to the contradiction that S itself contains a primitive skew curve, namely, the curve $T(\Pi)$. Having disposed of Case 1, we turn to

Case 2: $\Pi - \mathbf{T}$ contains an *infinite* number of distinct components.

Then if $e_1, e_2, \cdots, e_n, \cdots$ is the set of components of $\Pi - \mathbf{T}$, it is easily seen that $\lim_{n=\infty} \delta(e_n) = 0$. Consider the sequence of continua

$$(1) \qquad C_1, C_2, \cdots, C_n, \cdots,$$

where $C_n = \mathbf{T} + \sum_{i=1}^{n} e_i$. Since $\Pi + \mathbf{T} = \mathbf{T} + \sum_{i=1}^{\infty} e_i$ is *not* a continuum of class $\mathfrak{R}$, we conclude by Theorem E_1 that there is some continuum C_k of 1) that *fails* to belong to class $\mathfrak{R}$, and hence contains a *primitive skew curve* Π'. It is now quite evident that the point-set $\Pi' - \mathbf{T}$ contains but a *finite* number of components; consequently Case 2 reduces to Case 1, and our demonstration is completed.

V. The Immersion Theorems

THEOREM I_1. *The continuum* $\mathbf{M}$ *is homeomorphic with a subset of a peanian continuum of which the proper cyclic elements are simple closed surfaces.*

Demonstration. Consider the Hilbert[20] space $H^\omega = \{(x_1, x_2, \cdots, x_n, \cdots)\}$, and its subspaces

$$(1) \qquad H_1, H_2, \cdots, H_n, \cdots,$$

where H_n is defined in H^ω by the equation

$$(2) \qquad x_{2n} = x_{2(n+1)} = \cdots = x_{2(n+r)} = \cdots = 0.$$

The spaces H_n are each evidently homeomorphic with H^ω itself. Since the continuum $\mathbf{M}$ is, *a fortiori, compact* and *metric,* it follows from a theorem of Urysohn[21] that $\mathbf{M}$ is homeomorphic with a subset of H^ω; consequently $\mathbf{M}$ has a topological image M_1 in H_1.

We introduce a

Spanning Process: Given (1), a continuum M_n of class $\mathfrak{R}$ that lies in H_n, and (2), a *boundary curve* $\mathbf{T}_n$ of M_n; then $\mathbf{T}_n$ is *spanned* with a closed 2-cell $E_n^{(2)}$ in a fashion that

(1) $M_n \cdot E_n^{(2)} = \mathbf{T}_n$,

(2) $M_{n+1}, = M_n + E_n^{(2)}$, is a continuum of class $\mathfrak{R}$ that lies in H_{n+1}, and

(3) $\delta(E_n^{(2)}) \leqq \delta(\mathbf{T}_n) + \dfrac{1}{n}$.

---

[20] F. Hansdorff, Mengenlehre (Berlin, 1927), p. 98.

[21] Über die Metrisation der kompakten topologischen Räume, Math. Ann. 92 (1924), 275–293. Zum Metrisationsproblem, Math. Ann. 94 (1924), 309–315.

The 2-cell $E_n^{(2)}$ is determined as follows: Take an arbitrary point

(4) $p_n, = (x_1^{(n)}, x_2^{(n)}, \cdots, x_{2n-1}^{(n)}, 0, x_{2n+1}^{(n)}, 0, x_{2n+3}^{(n)}, 0, \cdots)$ on $\mathbf{T}_n$, and denote by

(5) P_n the point $(x_1^{(n)}, x_2^{(n)}, \cdots, x_{2n-1}^{(n)}, \frac{1}{n}, x_{2n+1}^{(n)}, 0, x_{2n+3}^{(n)}, 0, \cdots)$ of H_{n+1}.

Then let Z_n be a variable point on $\mathbf{T}_n$, and put

$$E_n^{(2)} = \sum_{Z_n \subset \mathbf{T}_n} P_n Z_n, \tag{6}$$

where $P_n Z_n$ represents, for a given point Z_n, the uniquely determined (closed) *straight interval* in H_{n+1} that is bounded by the two points P_n and Z_n.[22] The point-set $E_n^{(2)}$ is evidently a closed 2-cell lying in H_{n+1} and having $\mathbf{T}_n$ as its *boundary curve* (in our sense as well as in the combinatorial sense). Consequently, since $\mathbf{T}_n$ is the common part of M_n and $E_n^{(2)}$, it follows, by Theorem E_2, that the continuum $M_{n+1} = M_n + E_n^{(2)}$ belongs to class $\mathfrak{R}$ (and is clearly $\subset H_{n+1}$). The relations (4), (5), (6) imply

$$\delta(E_n^{(2)}) = \delta(\mathbf{T}_n + P_n) \leqq \delta(\mathbf{T}_n) + \rho(\mathbf{T}_n, P_n) = \delta(\mathbf{T}_n) + \rho(p_n, P_n);$$

and on utilizing

$$\rho(p_n, P_n) = \frac{1}{n}$$

we obtain

$$\delta(E_n^{(2)}) \leqq \delta(\mathbf{T}_n) + \frac{1}{n}.$$

It is seen, therefore, that $E_n^{(2)}$ satisfies the requirements 1), 2), and 3).

Let us assume that the subcontinuum M_1 of H_1 has at least *one* boundary curve; for otherwise, by Theorem 1, our proposition is trivial. Then if d_1 denotes the upper bound of the diameters of the *boundary curves* of M_1, it follows, by Theorem 2, that some such curve, $\mathbf{T}_1$, is itself of diameter d_1. With reference then to M_1 and $\mathbf{T}_1$, apply the spanning process and obtain the continuum $M_2 = M_1 + E_1^{(2)}$. If M_2 also has *boundary curves* let $\mathbf{T}_2$ be one of them

---

[22] If $P = (x_1, x_2, \cdots, x_n, \cdots)$ and $Q = (y_1, y_2, \cdots, y_n, \cdots)$ are an arbitrary pair of distinct points of H^ω, then we define the *straight interval* PQ to be the set of all points $(z_1, z_2, \cdots, z_n, \cdots)$ of H^ω whose coordinates satisfy the relation

$$d = \frac{z_1 - x_1}{y_1 - x_1} = \frac{z_2 - x_2}{y_2 - x_2} = \cdots = \frac{z_n - x_n}{y_n - x_n} = \cdots,$$

where $0 \leqq d \leqq 1$.

of maximal diameter d_2 and *span* $\mathbf{T}_2$, obtaining the continuum $M_3 = M_2 + E_2^{(2)}$. By continuing in this manner it is possible that a continuum

$$M_r = M_1 + E_1^{(2)} + \cdots + E_{r-1}^{(2)}$$

will be reached that has *no boundary curves*; in which case our theorem is proved since (by Theorem 1) the proper cyclic elements of M_r are simple closed surfaces and M_r contains the topological image M_1 of $\mathbf{M}$. If, however, on continuing the *spanning process* indefinitely, all the continua M_r, $r \geqq 1$, have *boundary curves*; then consider the sequence of continua

$$(7) \qquad M_1, E_1^{(2)}, E_2^{(2)}, \cdots, E_n^{(2)}, \cdots .$$

It has already been observed that

$$(8) \qquad M_n = M_1 + \sum_{i=1}^{n-1} E_i^{(2)} \text{ is a continuum of class } \mathfrak{R};$$

furthermore it is clear that

$$(9) \qquad M_1 \cdot E_n^{(2)} \neq 0, \qquad n = 1, 2, 3, \cdots;$$

and we now proceed to establish

$$(10) \qquad \operatorname*{Lim}_{n=\infty} \delta(E_n^{(2)}) = 0.$$

Suppose that (10) does not hold; then from (3) it follows that

$$\operatorname*{Lim}_{n=\infty} \delta(\mathbf{T}_n) \neq 0.$$

Consequently there exists an increasing sequence of positive integers

$$n_1, n_2, \cdots, n_i, \cdots$$

for which the sequence of *boundary curves*

$$(11) \qquad \mathbf{T}_{n_1}, \mathbf{T}_{n_2}, \cdots, \mathbf{T}_{n_i}, \cdots$$

converges to a sequential limit which is of diameter >0 and is hence, by Theorem 2, a boundary curve $\mathbf{T}$ of M_1.[23] Observe that the sequence (11) satisfies the hypothesis of Theorem 3; therefore three integers $u < v < w$ can be chosen sufficiently large that the curves $\mathbf{T}_{n_u}$, $\mathbf{T}_{n_v}$, $\mathbf{T}_{n_w}$ will assuredly have at least *four* distinct common points a, b, c, d which we assume to lie on $\mathbf{T}_{n_w}$ in the order $abcd$. Now consider the closed 2-cells $E_{n_u}^{(2)}$ and $E_{n_v}^{(2)}$ that *span* the curves $\mathbf{T}_{n_u}$ and $\mathbf{T}_{n_v}$ respectively. In $E_{n_u}^{(2)}$ take an arc ac having only its end points a and c on $\mathbf{T}_{n_u}$; and similarly in $E_{n_v}^{(2)}$ take an arc bd having only b and d on $\mathbf{T}_{n_v}$. The arcs ac and bd evidently have no points in common, and, since the point pairs (a, c) and (b, d) separate each other on $\mathbf{T}_{n_w}$, it follows, from Lemma 1, that $\mathbf{T}_{n_w}$

[23] That $\mathbf{T}$ is a subset of M_1 follows from $\mathbf{T}_n \subset M_1$ for all values of n.

is not a boundary curve of the continuum M_{n_w}, for the relation $n_w > n_v > n_u$ implies $M_{n_w} \supset E^{(2)}_{n_u} + E^{(2)}_{n_v} \supset ac + bd$. In view of this contradiction, the validity of relation (10) is established.

From (8), (9), and (10) it is now apparent that the sequence of continua (7) satisfies the hypothesis of Theorem E_1; consequently the point set

$$M_\omega = M_1 + \sum_{n=1}^{\infty} E^{(2)}_n$$

is a continuum of class $\mathfrak{R}$. Furthermore, M_ω does *not* have a *boundary curve* for if $\mathbf{T}$ were such a curve, then $\lim_{n=\infty} \delta(\mathbf{T}_n) = 0$ insures $\delta(\mathbf{T}) > \delta(\mathbf{T}_n)$ for n sufficiently large, implying the contradiction that $\mathbf{T}_n$ is not one of the *boundary curves* of maximal diameter d_n of M_n. We conclude, then, from Theorem 1, that the proper cyclic elements of M_ω are simple closed surfaces. Our proposition is therefore proved, since M_ω contains the continuum M_1 which is homeomorphic with $\mathbf{M}$.

The preceding theorem has as an immediate consequence the following corollaries:

COROLLARY 1. *A cyclic continuum of class $\mathfrak{R}$ that has precisely one boundary curve is a closed 2-cell.*

This corollary, together with its apparent converse, constitutes another characterization of the closed 2-cell; others having recently been given, independently, by D. W. Woodard,[24] H. Whitney,[25] and L. Zippin.[26]

COROLLARY 2. *Every continuum $\mathbf{M}$ of class $\mathfrak{R}$ is of Urysohn-Menger dimension $\leqq 2$.*

For, the cyclic elements of $\mathbf{M}$ are, by Theorem I_1, each homeomorphic with a subset of a spherical surface, and hence of dimension $\leqq 2$. Consequently $\mathbf{M}$ itself is of dimension $\leqq 2$.[27]

Let us now consider Theorem I_1 with reference only to the *cyclic* continua of class $\mathfrak{R}$; and observe, then, the interesting result that they are all homeomorphic with a subset of a spherical surface. We have consequently established the sufficiency[28] aspect the following interesting

THEOREM: *A necessary and sufficient condition that a cyclic peanian continuum be homeomorphic with a subset of a spherical surface is that it does not contain a primitive skew curve.*

That this theorem does not hold for peanian continua in general is evident when one considers a point-set consisting of a closed 2-cell $E^{(2)}$ and an arc B

---

[24] Bull. Amer. Math. Soc., vol. 36 (1930), p. 355.

[25] Tran. Amer. Math. Soc., vol. 35 (1933), pp. 261–273.

[26] Amer. Jour. Math., vol. 55 (1933), pp. 207–217.

[27] C. Kuratowski, Quelques applications d'éléments cycliques de M. Whyburn. Fund. Math., vol. 14 (1929), Theorem III, p. 142.

[28] The necessity proof is trivial.

such that $B \cdot E^{(2)}$ is an *interior* point of $E^{(2)}$. Accordingly, we proceed now to determine the *non-cyclic* continua of class $\mathfrak{R}$ that have topological images in the surface of a sphere.

THEOREM I_2. *If the continuum* $\mathbf{M}$ *satisfies the condition* (L)*: each cut-point* P *of* $\mathbf{M}$ *is a boundary point*[29] *of the closure of every component of* $\mathbf{M} - P$*; then* $\mathbf{M}$ *is homeomorphic with a subset of a spherical surface.*

Demonstration. Suppose that we are given an arbitrary positive number ϵ_n, a non-cyclic continuum M_n of class $\mathfrak{R}$, and a cyclic element C_n of M_n such that

α) M_n is a subset a euclidean six space R_6,

β) M_n satisfies the condition (L),

γ) C_n is *not* an end-point of M_n;

then, we introduce an

Augmentation Process for M_n relative to C_n and ϵ_n: Consider the collection $\mathbf{S}_n$ of *boundary arcs* $P_n^\sigma x_n^\sigma Q_n^\sigma$ of M_n, where $P_n^\sigma x_n^\sigma Q_n^\sigma$ is the sum of its subarcs $P_n^\sigma x_n^\sigma$ and $x_n^\sigma Q_n^\sigma$ in a fashion that

(1) x_n^σ is a cut-point of M_n that is contained in C_n,

(2) $\delta(P_n^\sigma x_n^\sigma) = \delta(x_n^\sigma Q_n^\sigma)$, and

(3) M_n is the sum of two continua X_n^σ and Y_n^σ for which

(i) $X_n^\sigma \supset C_n + P_n^\sigma x_n^\sigma, \quad Y_n^\sigma \supset x_n^\sigma Q_n^\sigma;$ and

(ii) $X_n^\sigma \cdot Y_n^\sigma = x_n^\sigma.$

The condition (L) insures that the collection $\mathbf{S}_n$ is non-vacuous. Let $\bar{u}_n$ denote the upper bound of the diameters of the arcs of $\mathbf{S}_n$ and take $P_n x_n Q_n$ to be one of them for which

$$\delta(P_n x_n Q_n) > \bar{u}_n/2. \tag{4}$$

Since M_n is, by Corollary 2, of dimension $\leqq 2$, and consequently, according to the Menger Immersion Theorem,[30] homeomorphic with a subset of euclidean five-space R_5, it follows from a theorem due to Mazurkewicz[31] that M_n is accessible at each of its points from the point-set $R_6 - M_n$. Hence (since M_n

[29] Cf. Introduction: Definition II.

[30] This theorem, which was enunciated by Menger, is to the effect that every compact n-dimensional space is homeomorphic with a subset of R_{2n+1}. Cf. K. Menger, Dimensionstheorie (Leipzig), 1928, pp. 287–303. Various proofs of this theorem have been given: see papers by G. Nobeling, Math. Annalen, vol. 104 (1930); L. Pontrjagin and G. Tolstowa, Math. Annalen, vol. 105 (1931); and S. Lefschetz, Annals of Math. (2), vol. 32 (1931).

[31] Sur un problem of M. Knaster, Fund. Math., XIII, pp. 146–150.

does not *separate* any region of R_6) there exists an arc $P_n y_n Q_n$ joining the point P_n to the point Q_n such that

$$\widehat{P_n y_n Q_n} \subset R_6 - M_n, \tag{5}$$

and

$$\delta(P_n y_n Q_n) < \rho(P_n, Q_n) + \epsilon_n. \tag{6}$$

Let us now set

$$M_{n+1} = M_n + P_n y_n Q_n. \tag{7}$$

The point-set M_{n+1} is called an *augmentation set* of M_n relative to C_n and ϵ_n; it is not difficult to show that M_{n+1} is a continuum of class $\mathfrak{R}$ which satisfies the relations α), β), and γ).

Choose a sequence of positive numbers $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_n > \cdots$ that converges to zero. Then take in R_6 a topological image M_1 of $\mathbf{M}$, and let C_1 be an arbitrary cyclic element of M_1 that is *not* an end point. With reference to C_1 and ϵ_n, we now apply the augmentation process to M_1 and obtain $M_2 = M_1 + P_1 y_1 Q_1$. If M_2 is *not* cyclic, then let C_2 be *that* one of its cyclic elements which contains $P_1 y_1 Q_1$, and obtain next an augmentation set M_3 $(= M_2 + P_2 y_2 Q_2)$ of M_2 relative to C_2 and ϵ_2. In general, if M_n *fails* to be cyclic, let C_n be *that* one of its cyclic elements which contains the arc $P_{n-1} y_{n-1} Q_{n-1}$, and proceed to obtain the augmentation set M_{n+1} of M_n relative to C_n and ϵ_n.

If, for some value r of n, the continuum M_r is *cyclic*, then our theorem is proved. For $\mathbf{M}$ has a topological image M_1, in M_r; and by Theorem I_1, M_r (being a continuum of class $\mathfrak{R}$) is homeomorphic with a subset of a spherical surface. If, however, for all values of n the continuum M_n has a cut-point, then consider the infinite sequence of continua

$$M_1,\ P_1 y_1 Q_1,\ P_2 y_2 Q_2,\ \cdots,\ P_n y_n Q_n,\ \cdots.$$

We have already seen that the continuum $M_n = M_1 + \sum_{i=1}^{n-1} P_i y_i Q_i$ belongs to class $\mathfrak{R}$. It is, furthermore, easy to show

$$M_1 \cdot (P_n y_n Q_n) \neq 0 \qquad n = 1, 2, \cdots,$$

for as a matter of fact the relations (3), (7), and

$$C_1 \subset C_2 \subset C_3 \cdots \subset C_n \subset \cdots$$

readily imply

$$M_1 \cdot (P_n y_n Q_n) \supset Q_n,$$

for all values of n. Finally, by a careful consideration of the manner in which the arcs $P_n y_n Q_n$ were chosen (with particular reference to relation (6)), it can be shown that

$$\lim_{n=\infty} \delta(P_n y_n Q_n) = 0.$$

Therefore, as a result of Theorem E_1, we conclude that the point-set

$$M_\omega = M_1 + \sum_{n=1}^{\infty} P_n y_n Q_n$$

is a continuum of class $\mathfrak{N}$ which, in virtue of relation (4), must also be *cyclic*. Our theorem is consequently established, since $M_1 \subset M_\omega$ and M_ω is homeomorphic with a subset of a spherical surface.

Scarcely any proof is needed in demonstrating the converse of Theorem I_2; and we have now obtained a solution to our original problem that may be stated as follows:

A necessary and sufficient condition that a peanian continuum K be homeomorphic with a subset of a spherical surface is that

(1) *K does not contain a primitive skew curve, and*

(2) *each cut-point P of K is a boundary point of the closure of every component of $K - P$.*

THE UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PA.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

ANALOGUES OF THE JACOBI CONDITION FOR THE PROBLEM OF MAYER IN THE CALCULUS OF VARIATIONS[1]

By William T. Reid

(Received July 8, 1933)

1. **Introduction.** The general problem of Mayer in the calculus of variations as formulated by Bliss [I][2] is that of finding in a class of arcs

$$y_i = y_i(x) \qquad (i = 1, \cdots, n; x_1 \leqq x \leqq x_2), \tag{1}$$

satisfying the differential equations and end-conditions

$$\varphi_\alpha(x, y, y') = 0 \qquad (\alpha = 1, \cdots, m < n), \tag{2}$$

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 1), \tag{3}$$

one which minimizes a function of the form

$$g[x_1, y(x_1), x_2, y(x_2)]. \tag{4}$$

This problem has been shown by Bliss [I, pp. 313, 314] to be equivalent to the problem of Bolza in the sense that each can be transformed into one of the other type. In the problem of Bolza the expression to be minimized is of the form

$$I = G[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f(x, y, y')\, dx.$$

For the problem of Mayer defined above, Bliss [I; see also IV] deduced the differential equations and transversality conditions that must be satisfied by a minimizing arc. For this problem Cope [II] has given a further necessary condition for a minimum, corresponding to that of Jacobi for simpler problems, in terms of the characteristic numbers of a boundary value problem associated with the second variation.[3] For the problem of Bolza, Morse has stated an analogue of Jacobi's condition in terms of the characteristic numbers of a boundary value problem, and has shown that the necessary conditions which he lists, when suitably strengthened, are sufficient for a minimum [V]. More recently Bliss [VII] has also established sufficient conditions for the problem of Bolza. A feature of the method used by Bliss is that the analogue of Jacobi's condition is phrased in terms of the characteristic numbers of a quadratic form involving only a finite number of variables, instead of in terms of the character-

[1] Presented to the American Mathematical Society, Dec. 30, 1932; §6 of this paper was presented separately under the title "Note on a preceding paper."

[2] The numbers in the square brackets refer to the bibliography at the end of this paper.

[3] For references to earlier literature on the boundary value problem method of attack, see VII, p. 261.

istic numbers of a boundary value problem associated with the accessory differential equations.

In treating the problem of Bolza hypotheses were made by Morse and Bliss which implied that the function f was not identically zero, and the sets of sufficient conditions established by them were, therefore, not applicable directly to the problem of Mayer. Recently sufficient conditions for the problem of Mayer with variable end-points have been established by Bliss and Hestenes [VIII] and Hestenes [IX]. For this problem also the analogue of Jacobi's condition has been given in the same form as that introduced by Bliss for the problem of Bolza [see IX, §4], and the condition therefore differs in form from that originally given by Cope.

It is the purpose of the present paper to show the connection between these two different phrasings of the analogues of Jacobi's condition for the problem of Mayer. In §2 there is given the boundary value problem formulation of Jacobi's condition and in §3 there is stated the form of this condition used by Bliss. The relations between these two different formulations are considered in §4, and in §5 it is shown that whenever the end-points x_1 and x_2 are not conjugate the analogue of Jacobi's condition that Bliss has given may be expressed in terms of the characteristic numbers of a boundary value problem of a different nature than the problem considered by Cope. In connection with the formulation of the condition as given by Cope use will be made of some results on boundary value problems recently established by the author [X], and of some extension of these results obtained in §6 of the present paper.

A particular arc E_{12} whose minimizing properties are to be studied is supposed to satisfy the hypotheses that are customarily made [see VIII, §2 and IX, §1]. The first necessary condition for a minimum, the Weierstrass condition, and the condition of Clebsch will be designated, respectively, by the numerals I, II and III. For a precise statement of these conditions and notations used throughout this paper the reader is referred to the above indicated papers of Bliss and Hestenes [III, VII, VIII and IX].[4]

2. **The boundary value problem formulation of Jacobi's condition.** Consider a normal non-singular admissible arc E_{12} without corners, with end-values satisfying the equations $\psi_\mu = 0$, and which satisfies the first necessary condition I. Such an arc is necessarily an extremal [III, p. 684]. Let $\xi_1, \xi_2, \eta_i(x)$ be a set of admissible variations along E_{12} satisfying the equations

$$\Phi_\alpha(x, \eta, \eta') \equiv \varphi_{\alpha y'_i}\eta'_i + \varphi_{\alpha y_i}\eta_i = 0, \tag{5}$$

$$\begin{aligned}\Psi_\mu[\xi_1, \xi_2, \eta(x_1), \eta(x_2)] \equiv (\psi_{\mu x_1} + \psi_{\mu y_{i1}}y'_{i1})\,\xi_1 + (\psi_{\mu x_2} + \psi_{\mu y_{i2}}y'_{i2})\xi_2 \\ + \psi_{\mu y_{i1}}\eta_i(x_1) + \psi_{\mu y_{i2}}\eta_i(x_2) = 0\,.\end{aligned} \tag{6}$$

[4] Graves has recently proven the necessity of conditions II and III for minimizing arcs which are assumed to be only normal, in place of the assumption of normality on every sub-interval used by Bliss. See Graves, On the Weierstrass condition for the problem of Bolza in the calculus of variations, Annals of Mathematics, Vol. 33 (1932), pp. 747–753.

There is then a one-parameter family of admissible arcs

$$y_i = y_i(x, b) \qquad x_1(b) \leqq x \leqq x_2(b)$$

satisfying the end-conditions (3), containing E_{12} for $b = 0$, and having ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12} [III, pp. 678, 695]. The functions $x_1(b)$, $x_2(b)$, $y_i(x, b)$ and $y_{ib}(x, b)$ are continuous in a neighborhood of the values (x, b) defining E_{12}, and their derivatives x_{1b}, x_{2b}, x_{1bb}, x_{2bb}, y_{ix}, y_{ixbb} and y_{ibb} have the same property except possibly at the values of x defining the corners of the arc $\eta_i = \eta_i(x)$ in $x\eta$-space.

The second variation g_2 along E_{12} is then expressible in the form

$$g_2[\xi_1, \xi_2, \eta] = 2H[\xi_1, \xi_2, \eta(x_1), \eta(x_2)] + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\, dx\,,$$

where H is a quadratic form in its arguments, and

$$2\omega = F_{y_i' y_j'}\eta_i'\eta_j' + 2F_{y_i' y_j}\eta_i'\eta_j + F_{y_i y_j}\eta_i\eta_j$$

[see II, §2]. Here, as elsewhere, F is the function $\lambda_\alpha(x)\varphi_\alpha$ determined by the first necessary condition. Let

$$\text{(7)} \qquad \Omega(x, \eta, \eta', \mu) = \omega(x, \eta, \eta') + \mu_\alpha \Phi_\alpha(x, \eta, \eta'),$$

$$\text{(8)} \qquad W(x, \xi, \eta, \xi', \eta', \mu) = \Omega(x, \eta, \eta', \mu) + \mu_{m+1}\xi_1' + \mu_{m+2}\xi_2'\,,$$

and consider the differential system consisting of the equations

$$\text{(9)} \quad \begin{aligned} &\text{a)} \quad \frac{d}{dx} W_{\eta_i'} - W_{\eta_i} + \lambda\eta_i \equiv \frac{d}{dx}\Omega_{\eta_i'} - \Omega_{\eta_i} + \lambda\eta_i = 0, \qquad \mu_{m+1}' = 0 = \mu_{m+2}'\,, \\ &\text{b)} \quad \Phi_\alpha(x, \eta, \eta') = 0, \qquad \xi_1' = 0 = \xi_2'\,, \end{aligned}$$

together with the two-point boundary conditions

$$\text{(10a)} \quad \begin{aligned} H_{\eta_i 1} + d_\mu\Psi_{\mu\eta_i 1} - \Omega_{\eta_i'}(x, \eta, \eta', \mu)\,\Big|^{x=x_1} &= 0\,, \\ H_{\eta_i 2} + d_\mu\Psi_{\mu\eta_i 2} + \Omega_{\eta_i'}(x, \eta, \eta', \mu)\,\Big|^{x=x_2} &= 0\,, \\ H_{\xi_1} + d_\mu\Psi_{\mu\xi_1} - \lambda\xi_1(x_1) - \mu_{m+1}(x_1) &= 0\,, \\ H_{\xi_2} + d_\mu\Psi_{\mu\xi_2} - \lambda\xi_2(x_2) + \mu_{m+2}(x_2) &= 0\,, \\ \mu_{m+2}(x_1) &= 0\,, \\ \mu_{m+1}(x_2) &= 0\,, \end{aligned}$$

$$\text{(10b)} \qquad \Psi_\mu[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)] = 0\,.$$

In the equations (10a) the partial derivatives of H are supposed to have the arguments $[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)]$.

A value λ will be said to be a characteristic number of this differential system if for this value there exist functions $\eta_i(x)$, $\xi_1(x)$, $\xi_2(x)$ of class C'' with multipliers $\mu_\alpha(x)$, $\mu_{m+1}(x)$, $\mu_{m+2}(x)$ of class C' such that the set η_i, ξ_1, ξ_2, μ_α, μ_{m+1}, μ_{m+2} does not vanish identically on x_1x_2, the set satisfies (9) on this interval,

and is such that there exist constants d_μ with which the end-values of the set satisfy (10). From the fact that E_{12} is assumed to be normal it follows that the only solution of the system for which $\eta_i \equiv 0 \equiv \xi_1 \equiv \xi_2$ on x_1x_2 is the identically vanishing solution $\eta_i \equiv 0 \equiv \xi_1 \equiv \xi_2 \equiv \mu_\alpha \equiv \mu_{m+1} \equiv \mu_{m+2}$.

Cope [II] has established the following analogue of the necessary condition of Jacobi, which will be denoted by IV_C.

IV_C. *If E_{12} is a normal non-singular minimizing arc without corners, then the inequality $\lambda \geqq 0$ must be satisfied by every characteristic number λ of the boundary value problem* (9), (10).

The boundary value problem (9), (10) is seen to consist of the Euler-Lagrange equations and transversality conditions for the problem of minimizing $g_2[\xi_1(x_1), \xi_2(x_2), \eta]$ in the class of admissible sets $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ which satisfy the differential equations and end-conditions (9b) and (10b), and which give a fixed value to the expression

$$(11) \qquad \xi_1^2(x_1) + \xi_2^2(x_2) + \int_{x_1}^{x_2} \eta_i(x)\, \eta_i(x)\, dx\ .$$

This boundary value problem is then of the form recently considered by the author [X], and if E_{12} is a normal extremal arc whose end-values satisfy the equations $\psi_\mu = 0$, and which satisfies conditions I and III′, then hypotheses ($H1$), ($H2$) and ($H3$) of that paper are seen to be satisfied; condition ($H4$) will not in general be satisfied, but by a linear change of the parameter λ the system (9), (10) may be reduced to a similar one for which hypothesis ($H4$) is satisfied.[5] If E_{12} is assumed to be normal on every sub-interval $x_1'x_2'$ of x_1x_2 then hypothesis ($H5$) of [X] will also be satisfied by the system (9), (10), but without this additional hypothesis the above defined boundary value problem may be shown to have infinitely many characteristic numbers, each of which is characterized by a corresponding minimizing property. This result is established in §6 of the present paper, and the proof depends upon the fact that if $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ is an admissible set which satisfies conditions (9b) and (10b), and is such that the expression (11) has the value zero, then $\xi_1(x) \equiv 0 \equiv \xi_2(x) \equiv \eta_i(x)$ on x_1x_2. In particular, the smallest characteristic number λ_1 is the minimum value of $g_2[\xi_1(x_1), \xi_2(x_2), \eta]$ in the class of admissible sets $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ which satisfy (9b) and (10b) and are normed such that the expression (11) has the value unity.[6] We have, therefore, the following result, which is much stronger than the above condition IV_C.

---

[5] Hypotheses ($H2$) and ($H3$) of [X] are, respectively, the assumptions that E_{12} be non-singular and normal; if E_{12} is then assumed to satisfy condition III it may be shown that (9), (10) can be reduced to a similar system for which ($H4$) is also satisfied. It has been pointed out to the author by Prof. Bliss, however, that the assumption that E_{12} be non-singular and satisfy condition III is equivalent to the assumption that E_{12} satisfy condition III′.

[6] This particular boundary value problem has also been treated by Hu in a manner quite different from that used in the author's paper [X]; see Hu, The problem of Bolza and its accessory boundary value problem, *Dissertation*, Chicago (1931).

THEOREM 2.1. *If E_{12} is a normal admissible arc without corners whose end-values satisfy the equations $\psi_\mu = 0$, and which satisfies conditions I and III', then a necessary and sufficient condition that the second variation $g_2[\xi_1, \xi_2, \eta]$ along E_{12} be non-negative for all sets of admissible variations $\xi_1, \xi_2, \eta_i(x)$ which satisfy equations* (6) *is that the smallest characteristic number λ_1 of the associated boundary value problem* (9), (10) *satisfy the relation $\lambda_1 \geqq 0$.*

3. **Formulation of the Jacobi condition as given by Bliss.** As indicated in §1, for the problem of Bolza an analogue of Jacobi's condition has been stated by Bliss in terms of a quadratic form involving only a finite number of variables. The same formulation of the condition has also been used by Hestenes for the problem of Mayer [IX, §4]. If E_{12} is a normal non-singular extremal arc it is a member of a $(2n - 1)$-parameter family of extremals

$$(12)\qquad \begin{aligned} y_i &= y_i(x, c_1, \cdots, c_{2n-1}) = y_i(x, c) \qquad (x_1 \leqq x \leqq x_2), \\ \lambda_\alpha &= \lambda_\alpha(x, c_1, \cdots, c_{2n-1}) = \lambda_\alpha(x, c) \end{aligned}$$

for special values $(x_1, x_2, c) = (x_{10}, x_{20}, c_0)$. The functions

$$y_i, y_{ix}, z_i = F_{y'_i}(x, y, y', \lambda), z_{ix}, \lambda_\alpha$$

have continuous first and second partial derivatives in a neighborhood of the values (x, c) defining E_{12}, and for the special values

$$c_s = c_{s0} \qquad (s = 1, \cdots, 2n - 1)$$

the matrix

$$(13)\qquad \left\| \begin{matrix} y_{ic_s} & 0 \\ z_{ic_s} & z_i \end{matrix} \right\|$$

is non-singular on $x_{10} \leqq x \leqq x_{20}$ [VIII, §§3, 4].

Using the notation introduced above, a part of the analogue of the necessary condition of Jacobi for the problem of Mayer as given by Hestenes may be stated as follows [IX, §4; for the corresponding condition for the problem of Bolza, see VII, p. 266]:

IV_B. *If E_{12} is a normal non-singular minimizing arc without corners, then $g_2[dx_1, dx_2, y_{c_s}\, dc_s] \geqq 0$ for all sets $(dx_1, dx_2, dc_s) \neq (0, 0, 0)$ which satisfy the equations*

$$(14)\qquad \Psi_\mu[dx_1, dx_2, y_{c_s}(x_1)\, dc_s, y_{c_s}(x_2)\, dc_s] = 0\,.$$

It is to be remarked that the necessary condition IV_B follows readily from the fact that if (dx_1, dx_2, dc_s) is a set which satisfies the equations (14), then $\xi_1 = dx_1$, $\xi_2 = dx_2$, $\eta_i(x) = y_{ic_s}(x) dc_s$ is a set of admissible variations such that $\Psi_\mu[\xi_1, \xi_2, \eta(x_1), \eta(x_2)] = 0$ and for which, therefore, $g_2[\xi_1, \xi_2, \eta] \geqq 0$.

Besides the condition IV_B, the fourth necessary condition as given by Hestenes includes the following condition, which I have chosen to list separately

since it is proved quite independently of IV_B and the condition is a consequence of the theory of the simpler type of the Mayer problem [VIII, §5].

IV_0. *If E_{12} is a normal non-singular minimizing arc without corners and normal on every sub-interval x_3x_2 $(x_1 < x_3 < x_2)$, then there can be no point conjugate to 1 on E_{12} between 1 and 2.*

4. **Relations between conditions IV_C, IV_B, and IV_0.** From the discussion in the preceding sections, the following result is immediate:

THEOREM 4.1. *If E_{12} is a normal admissible arc without corners whose end-values satisfy the equations $\psi_\mu = 0$, and which satisfies conditions I and III', then whenever condition IV_C is satisfied the condition IV_B is also satisfied; furthermore, if E_{12} is normal on every sub-interval x_3x_2 $(x_1 < x_3 < x_2)$ condition IV_0 is also satisfied.*

It may be shown by a very simple example that the converse of this theorem is not true. Let $n = 2$, $m = 1$, $p = 4$ and

$$g = y_2(x_2), \qquad \psi_1 = x_1, \qquad \psi_2 = x_2 - \pi, \qquad \psi_3 = y_1(x_1) - y_1(x_2),$$
$$\psi_4 = y_2(x_1), \qquad \varphi_1 = y_1'^2 - y_1^2 - y_2'.$$

For this problem the admissible arc

$$E_{12} : y_1(x) \equiv 0 \equiv y_2(x) \qquad (0 \leqq x \leqq \pi)$$

is normal, satisfies the first necessary condition I, and along it we have

$$
\begin{aligned}
&g_2[\xi_1, \xi_2, \eta] = \int_0^\pi [\eta_1'^2 - \eta_1^2]\, dx\,, \\
&\Psi_1 = \xi_1, \qquad \Psi_2 = \xi_2, \qquad \Psi_3 = \eta_1(0) - \eta_1(\pi), \qquad \Psi_4 = \eta_2(0)\,, \qquad (15)\\
&\Phi_1 = \eta_2'(x)\,.
\end{aligned}
$$

Along the arc E_{12} condition IV_0 is clearly satisfied; also, since the only solution of the differential equation $\eta_1'' + \eta_1 = 0$ which satisfies the condition $\Psi_3 = 0$ is of the form $\eta_1 = k \sin x$ and for such functions the integral (15) vanishes, the condition IV_B is seen to be satisfied. On the other hand, $\xi_1 \equiv \xi_2 \equiv \eta_2 \equiv 0$, $\eta_1 \equiv 1$ is an admissible set which satisfies $\Psi_\mu = 0$ $(\mu = 1, \cdots, 4)$, $\Phi_1 = 0$, and for which $g_2[\xi_1, \xi_2, \eta] = -\pi$. By Theorem 2.1 it then follows that condition IV_C is not satisfied; in fact, $\lambda = -1$ is seen to be a characteristic number of the corresponding boundary value problem.

Concerning conditions IV_C and IV_B we may prove, however, the following result:

THEOREM 4.2. *Let E_{12} be a normal admissible arc without corners whose end-values satisfy the equations $\psi_\mu = 0$, which is normal on every sub-interval x_1x_3 $(x_1 < x_3 \leqq x_2)$, for which conditions I and III' are satisfied, and such that there is no point conjugate to 1 on $x_1 < x \leqq x_2$. Then, whenever condition IV_B is satisfied by E_{12} the condition IV_C is also satisfied.*

The arc E_{12} is an extremal and since it is normal on x_1x_2 and the point 2 is not conjugate to 1, the matrix

$$\left\|\begin{matrix} y_{ic_s}(x_1) \\ y_{ic_s}(x_2) \end{matrix}\right\|, \tag{16}$$

where the functions $y_{ic_s}(x)$ are those defined in §3, has rank $2n-1$ [VIII, §6, Theorem 6.1]. Furthermore, since E_{12} is normal on every sub-interval x_1x_3 and there is no point conjugate to 1 on $x_1 < x \leqq x_2$, there exists a conjugate system of solutions U_{ik}, V_{ik} $(k = 1, \cdots, n)$ of the accessory equations in canonical form such that the determinant $|U_{ik}(x)|$ is different from zero on the entire interval x_1x_2 [VIII, §8].

For an arbitrary set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ along E_{12} the functions $\eta_i(x)$ are seen to satisfy the equations

$$F_{y_i'}(x_2)\eta_i(x_2) - F_{y_i'}(x_1)\eta_i(x_1) = 0\,. \tag{17}$$

Equations (17) hold, in particular, for $\eta_i(x) = y_{ic_s}(x)$ $(s = 1, \cdots, 2n-1)$; therefore, since the functions $F_{y_i'}$ do not all vanish either at x_1 or x_2 and since the matrix (16) is of rank $2n-1$, it follows that if ξ_1, ξ_2, $\eta_i(x)$ is an arbitrary set of admissible variations there exist unique values $dx_1(\xi_1)$, $dx_2(\xi_2)$, $dc_s(\eta)$ such that $\xi_1 = dx_1(\xi_1)$, $\xi_2 = dx_2(\xi_2)$, $\eta_i(x_1) = y_{ic_s}(x_1)dc_s(\eta)$, $\eta_i(x_2) = y_{ic_s}(x_2)dc_s(\eta)$. As the determinant $|U_{ik}|$ for the conjugate system U_{ik}, V_{ik} does not vanish on x_1x_2, it is a consequence of the theory of the Lagrange problem with fixed end-points that

$$g_2[\xi_1, \xi_2, \eta] \geqq g_2[dx_1(\xi_1), dx_2(\xi_2), y_{c_s}dc_s(\eta)]^{7}\,. \tag{18}$$

Since E_{12} satisfies condition IV_B we have

$$g_2[dx_1, dx_2, y_{c_s}dc_s] \geqq 0 \tag{19}$$

for all sets (dx_1, dx_2, dc_s) which satisfy the equations (14), and in view of the results of §2 it then follows that condition IV_C is also satisfied by E_{12}.

The notations IV_C' and IV_B' will be used to designate the conditions IV_C and IV_B when strengthened to exclude the equality signs which occur in their statements. If the smallest characteristic number of the boundary value problem (9), (10) is positive, it is a consequence of the minimizing property of this characteristic number that the second variation (7) is positive for all sets of admissible variations ξ_1, ξ_2, $\eta_i(x)$ that satisfy equations (6) and are not identically zero; in particular, $g_2[dx_1, dx_2, y_{c_s}dc_s]$ is positive for all sets

$$(dx_1, dx_2, dc_s) \neq (0, 0, 0)$$

which satisfy the equations (14), and therefore condition IV_B' is satisfied. If

[7] See III, pp. 737–739; also Bliss, The transformation of Clebsch in the calculus of variations, Proceedings of the International Mathematical Congress Held in Toronto, Vol. 1 (1924), pp. 589–603.

condition IV'_C is satisfied it is also readily seen that there is no point conjugate to 1 on $x_1 < x \leqq x_2$. We have, therefore

THEOREM 4.3. *If E_{12} is a normal admissible arc without corners whose end-values satisfy the equation $\psi_\mu = 0$ and which satisfies conditions I and III′, then whenever condition IV'_C is satisfied the condition IV'_B is also satisfied and there is no point conjugate to 1 on $x_1 < x \leqq x_2$.*

If E_{12} is a normal extremal arc normal on every sub-interval x_1x_3 ($x_1 < x_3 \leqq x_2$), which satisfies conditions I and III′, and is such that there is no point conjugate to 1 on $x_1 < x \leqq x_2$, then the equality sign in (18) holds only if

$$\eta_i(x) \equiv y_{ic_s}(x)\, dc_s(\eta) \quad (x_1 \leqq x \leqq x_2)\,.$$

If E_{12} also satisfies IV'_B, then the equality sign in (19) holds only if

$$(dx_1, dx_2, dc_s) = (0, 0, 0)\,.$$

It is also to be remarked that if E_{12} satisfies IV'_B then the point 2 can not be conjugate to 1. These results are combined in the following theorem:

THEOREM 4.4. *If E_{12} is a normal admissible arc without corners normal on every sub-interval x_1x_3 ($x_1 < x_3 \leqq x_2$) which satisfies conditions I, III′, IV_0, IV'_B, then the condition IV'_C is also satisfied on E_{12}.*

5. **Further discussion of condition $\mathrm{IV_B}$.** In this section it will be shown that whenever the ends of the extremal arc E_{12} are not conjugate condition IV_B may be expressed in terms of a boundary value problem of a different type from that discussed in §2. Using the notation of §3, we have for each set of values dx_1, dx_2, dc_s ($s = 1, 2, \cdots, 2n-1$) that

$$\begin{aligned} g_2[dx_1, dx_2, y_{c_s}dc_s] &= 2H[dx_1, dx_2, y_{c_s}(x_1)dc_s, y_{c_s}(x_2)dc_s] \\ &\quad + dc_s y_{ic_s}(x)\, \Omega_{\eta'_i}[x, y_{c_t}dc_t, y'_{c_t}dc_t, \lambda_{c_t}dc_t]\Big|_{x=x_1}^{x=x_2}\,. \end{aligned} \tag{20}$$

For simplicity of notation, let $z_1 = dx_1$, $z_2 = dx_2$, $z_{2+s} = dc_s (s = 1, 2, \cdots, 2n-1)$, and denote the quadratic form (20) by

$$2A[z] = a_{\sigma\tau}z_\sigma z_\tau \qquad (\sigma, \tau = 1, \cdots, 2n+1)\,, \tag{21}$$

and the linear equations (14) by

$$B_\mu[z] = b_{\mu\tau}z_\tau = 0 \qquad (\mu = 1, 2, \cdots, p)\,. \tag{22}$$

If $k_{\sigma\tau}z_\sigma z_\tau$ is an arbitrary positive definite quadratic form, then the conditions under which $A[z] \geqq 0$ for the set of values satisfying $B_\mu[z] = 0$ are the same as those under which $A[z] \geqq 0$ for the set of values satisfying

$$B_\mu[z] = 0, \qquad k_{\sigma\tau}z_\sigma z_\tau = 1\,. \tag{23}$$

If the matrix $\| b_{\mu\tau} \|$ is of rank p, then the minimum value of $2A[z]$ in the class of

values z which satisfy the conditions (23) is equal to the smallest zero l_1 of the determinant[8]

$$D(l) = \begin{vmatrix} a_{\sigma\tau} - lk_{\sigma\tau} & b\mu\sigma \\ b\mu\tau & 0 \end{vmatrix} . \tag{24}$$

There are seen to be $2n + 1 - p$ roots $l = l_h$ $(h = 1, \cdots, 2n + 1 - p)$ of the equation $D(l) = 0$, and corresponding to $l = l_h$ there exist constants $z_{\tau h}$, $d_{\mu h}$ not all zero and such that

$$\begin{aligned} a_{\sigma\tau}z_{\tau h} - l_h k_{\sigma\tau} z_{\tau h} + b_{\mu\sigma} d_{\mu h} = 0 \quad &(\sigma = 1, \cdots, 2n + 1; \\ b_{\mu\tau} z_{\tau h} = 0 \quad &\mu = 1, \cdots, p; h = 1, \cdots, 2n - 1 + p) . \end{aligned} \tag{25}$$

If E_{12} is normal on x_1x_2 and its end-points are not conjugate, then the matrix (16) has rank $2n - 1$ and the matrix $\| b_{\mu\tau} \|$ defined above is seen to have rank p; furthermore, the quadratic form

$$z_1^2 + z_2^2 + [y_{ic_s}(x_1)\, y_{ic_t}(x_1) + y_{ic_s}(x_2)\, y_{ic_t}(x_2)]\, z_{2+s} z_{2+t} \tag{26}$$

is positive definite. Upon defining the arbitrary quadratic form $k_{\sigma\tau}z_\sigma z_\tau$ as equal to the form (26), the equations (25) become

$$\begin{aligned} &\Psi_{\mu\xi_1} z_{1h} + \Psi_{\mu\xi_2} z_{2h} + [\Psi_{\mu\eta_i 1}\, y_{ic_s}(x_1) + \Psi_{\mu\eta_i 2}\, y_{ic_s}(x_2)]\, z_{2+s} = 0 , \\ &H_{\xi_1} - l_h z_{1h} + d_{\mu h}\, \Psi_{\mu\xi_1} = 0 , \\ &H_{\xi_2} - l_h z_{2h} + d_{\mu h}\, \Psi_{\mu\xi_2} = 0 , \\ &y_{ic_s}(x_1) \{ H_{\eta_i 1} - \Omega_{\eta_i'} \,|^{x=x_1} + d_{\mu h}\Psi_{\mu\eta_i 1} - l_h y_{ic_s}(x_1) z_{sh} \} \\ + &y_{ic_s}(x_2) \{ H_{\eta_i 2} + \Omega_{\eta_i'} \,|^{x=x_2} + d_{\mu h}\, \Psi_{\mu\eta_i 2} - l_h y_{ic_s}(x_2)\, z_{sh} \} = 0 , \end{aligned} \tag{27}$$

where the arguments of the partial derivatives of H are $[z_{1h}, z_{2h}, y_{c_s}(x_1)\, z_{2+s,h}, y_{c_s}(x_2) z_{2+s,h}]$ and those of $\Omega_{\eta_i'}$ are $[x, y_{c_s}(x)\, z_{2+s,h}, y'_{c_s}(x)\, z_{2+s,h}, \lambda_{c_s}(x)\, z_{2+s,h}]$.

From relation (17) for $\eta_i = y_{ic_s}(x)$ $(s = 1, \cdots, 2n - 1)$ and the last $2n - 1$ equations of (27), it follows that associated with the set $z_{\tau h}$ $(h = 1, \cdots, 2n + 1 - p)$ there is a constant ρ_h such that

$$\begin{aligned} &H_{\eta_i 1} + d_{\mu h}\Psi_{\mu\eta_i 1} - l_h y_{ic_s}(x_1)\, dc_s - [\Omega_{\eta_i'} + \rho_h \lambda_\alpha \varphi_{\alpha y_i'}]^{x=x_1} = 0 , \\ &H_{\eta_i 2} + d_{\mu h}\Psi_{\mu\eta_i 2} - l_h y_{ic_s}(x_2)\, dc_s - [\Omega_{\eta_i'} + \rho_h \lambda_\alpha \varphi_{\alpha y_i'}]^{x=x_2} = 0 . \end{aligned} \tag{28}$$

If we set $\xi_{1h}(x) \equiv z_{1h}$, $\xi_{2h}(x) \equiv z_{2h}$, $\eta_{ih}(x) = y_{ic_s}(x)\, z_{2+s,h}$, $\mu_{\alpha h}(x) = \lambda_{\alpha c_s}(x)\, z_{2+s,h} + \rho_h \lambda_\alpha$, then the set ξ_{1h}, ξ_{2h}, $\eta_{ih}(x)$, $\mu_{\alpha h}(x)$ is a non-identically vanishing solution of the differential equations

---

[8] See VII, §7; also Hancock, *Theory of maxima and minima*, Ginn and Co., Boston (1917), pp. 103–114. If the matrix $\| b_{\mu\tau} \|$ is of rank $p - r$, then r of the equations (22) are a consequence of the remaining equations, and $D(l)$ is replaced by a determinant of order $2n - 1 + p - r$.

$$(29)\qquad \begin{aligned} &a)\ \frac{d}{dx}\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \\ &b)\ \Phi_\alpha(x, \eta, \eta') = 0 = \xi_1' = \xi_2', \end{aligned}$$

which satisfies with constants $d_\mu = d_{\mu h}$ and $l = l_h$ the end-conditions

$$(30a)\qquad \begin{cases} H_{\eta i1} + d_\mu \Psi_{\mu\eta i1} - l\eta_i(x_1) - \Omega_{\eta_i'}(x, \eta, \eta', \mu) \,\big|^{x=x_1} = 0\,, \\ H_{\eta i2} + d_\mu \Psi_{\mu\eta i2} - l\eta_i(x_2) + \Omega_{\eta_i'}(x, \eta, \eta', \mu) \,\big|^{x=x_2} = 0\,, \\ H_{\xi 1} + d_\mu \Psi_{\mu\xi 1} - l\xi_1(x_1) = 0\,, \\ H_{\xi 2} + d_\mu \Psi_{\mu\xi 2} - l\xi_2(x_2) = 0\,, \end{cases}$$

$$(30b)\qquad \Psi_\mu[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)] = 0\,.$$

In the above equations the partial derivatives of H are supposed to have the arguments $[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)]$.

The boundary value problem (29), (30) is seen to have as characteristic numbers the $2n + 1 - p$ zeros of the determinant $D(l)$. We have, therefore, the following result:

THEOREM 5.1. *Suppose that E_{12} is a non-singular extremal arc normal on x_1x_2 whose end-values satisfy the equations $\psi_\mu = 0$, and which satisfies condition I. If, in addition, the point 2 is not conjugate to the point 1, then a necessary and sufficient condition that $g_2[dx_1, dx_2, yc_s, dc_s] \geqq 0$ $[> 0]$ for all sets $(dx_1, dx_2, dc_s) \neq (0, 0, 0)$ which satisfy equations (14) is that the smallest characteristic number l_1 of the boundary value problem (29), (30) satisfy the relation $l_1 \geqq 0$ $[l_1 > 0]$.*

If, in addition to the assumptions of the above theorem, E_{12} is assumed to satisfy condition III′, to be normal on every subinterval x_1x_3 $(x_1 < x_3 \leqq x_2)$, and to have no point conjugate to 1 on $x_1 < x \leqq x_2$, then by a method analogous to that used in the proof of Theorem 4.2 it may be proven that each characteristic value l_h is characterized by a corresponding minimizing property, and, in particular, that l_1 is the minimum value of $g_2[\xi_1(x_1), \xi_2(x_2), \eta]$ in the class of admissible sets $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ which satisfy conditions (29b), (30b), and are normed such that

$$\xi_1^2(x_1) + \xi_2^2(x_2) + \eta_i(x_1)\,\eta_i(x_1) + \eta_i(x_2)\,\eta_i(x_2) = 1\,.$$

It is to be remarked that the point 2 may be conjugate to 1, and yet the matrix of coefficients of equations (14) be of rank p. In this case it is still true that a necessary and sufficient condition that $g_2[dx_1, dx_2, y_{c_s}dc_s] \geqq 0$ for all sets (dx_1, dx_2, dc_s) which satisfy equations (14) is that the smallest zero of the determinant (24) be non-negative; one is no longer assured, however, that the zeros of (24) are characteristic numbers of the boundary value problem (29), (30). To illustrate this fact, consider the example given in §4. In this example, although the end-points are conjugate, the coefficient matrix of the equations corresponding to (14) is of rank four, and the determinant (24) has the

values z which satisfy the conditions (23) is equal to the smallest zero l_1 of the determinant[8]

$$D(l) = \begin{vmatrix} a_{\sigma\tau} - lk_{\sigma\tau} & b_{\mu\sigma} \\ b_{\mu\tau} & 0 \end{vmatrix} . \tag{24}$$

There are seen to be $2n + 1 - p$ roots $l = l_h$ $(h = 1, \cdots, 2n + 1 - p)$ of the equation $D(l) = 0$, and corresponding to $l = l_h$ there exist constants $z_{\tau h}$, $d_{\mu h}$ not all zero and such that

$$\begin{aligned} (25) \quad & a_{\sigma\tau}z_{\tau h} - l_h k_{\sigma\tau}z_{\tau h} + b_{\mu\sigma}d_{\mu h} = 0 \quad (\sigma = 1, \cdots, 2n + 1 ; \\ & b_{\mu\tau}z_{\tau h} = 0 \quad \mu = 1, \cdots, p; h = 1, \cdots, 2n - 1 + p) . \end{aligned}$$

If E_{12} is normal on x_1x_2 and its end-points are not conjugate, then the matrix (16) has rank $2n - 1$ and the matrix $\| b_{\mu\tau} \|$ defined above is seen to have rank p; furthermore, the quadratic form

$$z_1^2 + z_2^2 + [y_{ic_s}(x_1)\, y_{ic_t}(x_1) + y_{ic_s}(x_2)\, y_{ic_t}(x_2)]\, z_{2+s}z_{2+t} \tag{26}$$

is positive definite. Upon defining the arbitrary quadratic form $k_{\sigma\tau}z_\sigma z_\tau$ as equal to the form (26), the equations (25) become

$$(27) \quad \begin{aligned} & \Psi_{\mu\xi_1}z_{1h} + \Psi_{\mu\xi_2}z_{2h} + [\Psi_{\mu\eta_i 1}\, y_{ic_s}(x_1) + \Psi_{\mu\eta_i 2}\, y_{ic_s}(x_2)]\, z_{2+s} = 0 , \\ & H_{\xi_1} - l_h z_{1h} + d_{\mu h}\, \Psi_{\mu\xi_1} = 0 , \\ & H_{\xi_2} - l_h z_{2h} + d_{\mu h}\, \Psi_{\mu\xi_2} = 0 , \\ & y_{ic_s}(x_1) \left\{ H_{\eta_i 1} - \Omega_{\eta_i'} \Big|^{x=x_1} + d_{\mu h}\Psi_{\mu\eta_i 1} - l_h y_{ic_s}(x_1) z_{sh} \right\} \\ & + y_{ic_s}(x_2) \left\{ H_{\eta_i 2} + \Omega_{\eta_i'} \Big|^{x=x_2} + d_{\mu h}\, \Psi_{\mu\eta_i 2} - l_h y_{ic_s}(x_2)\, z_{sh} \right\} = 0 , \end{aligned}$$

where the arguments of the partial derivatives of H are $[z_{1h}, z_{2h}, y_{c_s}(x_1)\, z_{2+s,h}, y_{c_s}(x_2)z_{2+s,h}]$ and those of $\Omega_{\eta_i'}$ are $[x, y_{c_s}(x)\, z_{2+s,h}, y'_{c_s}(x)\, z_{2+s,h}, \lambda_{c_s}(x)\, z_{2+s,h}]$.

From relation (17) for $\eta_i = y_{ic_s}(x)$ $(s = 1, \cdots, 2n - 1)$ and the last $2n - 1$ equations of (27), it follows that associated with the set $z_{\tau h}$ $(h = 1, \cdots, 2n + 1 - p)$ there is a constant ρ_h such that

$$(28) \quad \begin{aligned} & H_{\eta_i 1} + d_{\mu h}\Psi_{\mu\eta_i 1} - l_h y_{ic_s}(x_1)\, dc_s - [\Omega_{\eta_i'} + \rho_h\lambda_\alpha\varphi_{\alpha y_i'}]^{x=x_1} = 0 , \\ & H_{\eta_i 2} + d_{\mu h}\Psi_{\mu\eta_i 2} - l_h y_{ic_s}(x_2)\, dc_s - [\Omega_{\eta_i'} + \rho_h\lambda_\alpha\varphi_{\alpha y_i'}]^{x=x_2} = 0 . \end{aligned}$$

If we set $\xi_{1h}(x) \equiv z_{1h}$, $\xi_{2h}(x) \equiv z_{2h}$, $\eta_{ih}(x) = y_{ic_s}(x)\, z_{2+s,h}$, $\mu_{\alpha h}(x) = \lambda_{\alpha c_s}(x)\, z_{2+s,h} + \rho_h\lambda_\alpha$, then the set ξ_{1h}, ξ_{2h}, $\eta_{ih}(x)$, $\mu_{\alpha h}(x)$ is a non-identically vanishing solution of the differential equations

---

[8] See VII, §7; also Hancock, *Theory of maxima and minima*, Ginn and Co., Boston (1917), pp. 103–114. If the matrix $\| b_{\mu\tau} \|$ is of rank $p - r$, then r of the equations (22) are a consequence of the remaining equations, and $D(l)$ is replaced by a determinant of order $2n - 1 + p - r$.

$$
(29) \qquad \begin{aligned} &a) \ \frac{d}{dx}\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \\ &b) \ \Phi_\alpha(x, \eta, \eta') = 0 = \xi_1' = \xi_2', \end{aligned}
$$

which satisfies with constants $d_\mu = d_{\mu h}$ and $l = l_h$ the end-conditions

$$
(30a) \qquad \begin{cases} H_{\eta_i 1} + d_\mu \Psi_{\mu \eta_i 1} - l\eta_i(x_1) - \Omega_{\eta_i'}(x, \eta, \eta', \mu) \,\big|^{x=x_1} = 0\,, \\ H_{\eta_i 2} + d_\mu \Psi_{\mu \eta_i 2} - l\eta_i(x_2) + \Omega_{\eta_i'}(x, \eta, \eta', \mu) \,\big|^{x=x_2} = 0\,, \\ H_{\xi_1} + d_\mu \Psi_{\mu \xi_1} - l\xi_1(x_1) = 0\,, \\ H_{\xi_2} + d_\mu \Psi_{\mu \xi_2} - l\xi_2(x_2) = 0\,, \end{cases}
$$

$$
(30b) \qquad \Psi_\mu[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)] = 0\,.
$$

In the above equations the partial derivatives of H are supposed to have the arguments $[\xi_1(x_1), \xi_2(x_2), \eta(x_1), \eta(x_2)]$.

The boundary value problem (29), (30) is seen to have as characteristic numbers the $2n + 1 - p$ zeros of the determinant $D(l)$. We have, therefore, the following result:

THEOREM 5.1. *Suppose that E_{12} is a non-singular extremal arc normal on x_1x_2 whose end-values satisfy the equations $\psi_\mu = 0$, and which satisfies condition I. If, in addition, the point 2 is not conjugate to the point 1, then a necessary and sufficient condition that $g_2[dx_1, dx_2, yc_s, dc_s] \geqq 0$ $[> 0]$ for all sets $(dx_1, dx_2, dc_s) \neq (0, 0, 0)$ which satisfy equations (14) is that the smallest characteristic number l_1 of the boundary value problem (29), (30) satisfy the relation $l_1 \geqq 0$ $[l_1 > 0]$.*

If, in addition to the assumptions of the above theorem, E_{12} is assumed to satisfy condition III′, to be normal on every subinterval x_1x_3 $(x_1 < x_3 \leqq x_2)$, and to have no point conjugate to 1 on $x_1 < x \leqq x_2$, then by a method analogous to that used in the proof of Theorem 4.2 it may be proven that each characteristic value l_h is characterized by a corresponding minimizing property, and, in particular, that l_1 is the minimum value of $g_2[\xi_1(x_1), \xi_2(x_2), \eta]$ in the class of admissible sets $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ which satisfy conditions (29b), (30b), and are normed such that

$$
\xi_1^2(x_1) + \xi_2^2(x_2) + \eta_i(x_1)\,\eta_i(x_1) + \eta_i(x_2)\,\eta_i(x_2) = 1\,.
$$

It is to be remarked that the point 2 may be conjugate to 1, and yet the matrix of coefficients of equations (14) be of rank p. In this case it is still true that a necessary and sufficient condition that $g_2[dx_1, dx_2, y_{c_s}dc_s] \geqq 0$ for all sets (dx_1, dx_2, dc_s) which satisfy equations (14) is that the smallest zero of the determinant (24) be non-negative; one is no longer assured, however, that the zeros of (24) are characteristic numbers of the boundary value problem (29), (30). To illustrate this fact, consider the example given in §4. In this example, although the end-points are conjugate, the coefficient matrix of the equations corresponding to (14) is of rank four, and the determinant (24) has the

single zero $l = 0$. The boundary value problem (29), (30), however, may be shown to be equivalent to

$$\eta_1'' + \eta_1 = 0$$

$$\eta_1(0) - \eta_1(\pi) = 0$$

$$\eta_1'(\pi) - \eta_1'(0) - l[\eta_1(0) + \eta_1(\pi)] = 0\,,$$

and is seen to have no characteristic values l.

6. Extension of previous results on boundary value problems. As indicated in §2, the boundary value problem (9), (10) of that section is of the type for which the author has given sufficient conditions for the existence of infinitely many characteristic numbers. Hypotheses $(H5^+)$ and $(H5^-)$ of [X] are closely related to the assumption of normality on every sub-interval of the given interval x_1x_2. It is the purpose of this section to prove that the existence of infinitely many characteristic numbers is obtained if hypotheses $(H5^+)$ and $(H5^-)$ are replaced by an independent hypothesis *which is satisfied by the boundary value problem* (9), (10) *of* §2 *without making the assumption that* E_{12} *is normal on every sub-interval of* x_1x_2.

Throughout this section the notation and hypotheses of [X] *will be used.*

Hypothesis $(H5)$. *If* η *is an admissible arc which is not identically zero on* x_1x_2, *then the quantity*

$$2G[\eta] + \int_{x_1}^{x_2} \eta_i K_{ij}\eta_j\,dx \tag{31}$$

is different from zero.

We shall here prove the following theorem:

Theorem 6.1. *If hypotheses* $(H1)$, $(H2)$, $(H3)$, $(H4)$ *and* $(H5)$ *are satisfied, then the boundary value problem* (2.4), (2.5) *of* [X] *has infinitely many characteristic numbers.*

It is readily shown that hypothesis $(H3)$ of [X] is equivalent to the assumption that there exist p differentially admissible arcs $w_{i\nu}$ $(\nu = 1, \cdots, p)$ such that the determinant $|\Psi_\gamma[w_\nu(x_1), w_\nu(x_2)]|$ is different from zero [see III; in particular, p. 693].

Suppose that the first k of the successive classes $S_1, S_2, \cdots$ and the first l of the successive classes $S_{-1}, S_{-2}, \cdots$ [see X, §5] are not empty. In particular, we may have $k = 0 = l$. The greatest lower bound of $I[\eta]$ in the class $S_q(q = -l, \cdots, -1, 1, \cdots, k)$ is the absolute value of a characteristic number corresponding to which there are r_q $[0 < r_q \leqq 2n]$ linearly independent solutions of the system (2.4), (2.5). This set of characteristic solutions and characteristic numbers will be represented by the symbols

$$\eta_{is}(x),\ \mu_{\alpha s}(x),\ \lambda_s\ (s = 1, \cdots, r = r_{-l} + \cdots + r_{-1} + r_1 + \cdots + r_k)\,;$$

furthermore, these solutions will be supposed to be chosen orthonormal in the sense that

$$G[\eta_s; \eta_t] + \int_{x_1}^{x_2} \eta_{is} K_{ij} \eta_{jt}\, dx = \delta_{st}[\operatorname{sgn} \lambda_s] \qquad (s, t = 1, \cdots, r),$$

where $\delta_{st} = 0$ if $s \neq t$, $\delta_{ss} = 1$; $\operatorname{sgn} \lambda_s = 1$ if $\lambda_s > 0$, $\operatorname{sgn} \lambda_s = -1$ if $\lambda_s < 0$. It will now be shown that there exist admissible arcs η which are not identically zero and which satisfy the relations

$$G[\eta_s; \eta] + \int_{x_1}^{x_2} \eta_{is} K_{ij} \eta_j\, dx = 0 \qquad (s = 1, \cdots, r). \tag{32}$$

As usual, we may enlarge the differential system $\Phi_\alpha(x, \eta, \eta') = 0$ to have the form

$$\begin{aligned} &\Phi_\alpha(x, \eta, \eta') = 0 && (\alpha = 1, \cdots, m) \\ &\Phi_\tau(x, \eta, \eta') = \zeta_\tau(x) && (\tau = m + 1, \cdots, n), \end{aligned} \tag{33}$$

where the functions Φ_τ are homogeneous and linear in the variables η_i, η_i', having the coefficients $\Phi_{\tau\eta_i'}$ and $\Phi_{\tau\eta_i}$ of class C'' on x_1x_2, and such that the matrix $\| \Phi_{i\eta_j'} \|$ is non-singular on x_1x_2 [see I].

Let $\zeta_\tau(x)$ be arbitrary continuous functions which can not be expressed as linear combinations of the $p + r$ functions

$$\Phi_\tau(x, w_\nu, w_\nu'), \qquad \Phi_\tau(x, \eta_s, \eta_s') \qquad (\nu = 1, \cdots, p; s = 1, \cdots, r)$$

on the interval x_1x_2. If $u \equiv (u_i(x))$ is a solution of the differential system (33) corresponding to such values $\zeta_\tau(x)$ it then follows that the functions $u_i(x)$ can not be expressed as a linear combination of the $p + r$ functions $w_{i\nu}(x)$, $\eta_{is}(x)$ on x_1x_2. Since the determinant $|\Psi_\gamma[w_\nu(x_1), w_\nu(x_2)]|$ is different from zero there exist unique constants c_ν such that the differentially admissible arc $v_i = u_i + c_\nu w_{i\nu}$ satisfies the equations $\Psi_\gamma[v(x_1), v(x_2)] = 0$ $(\gamma = 1, \cdots, p)$. Now set

$$\eta_i(x) = v_i(x) - \sum_{s=1}^{r} \{G[\eta_s; v] + \int_{x_1}^{x_2} \eta_{is} K_{ij} v_j\, dx\}\, \eta_{is}(x).$$

The arc η is seen to be an admissible arc which is not identically zero and which satisfies the r relations (32). In view of hypothesis $(H5)$ it then follows that either the class S_{k+1} or the class S_{-l-1} is not empty, and by induction it is established that there are infinitely many of the classes $S_1, S_{-1}, S_{-2}, \cdots$ which are not empty. From Theorems 5.1 and 5.2 of [X] it then follows that the boundary value problem (2.4), (2.5) has infinitely many characteristic numbers, and Theorem 6.1 is therefore established.

If we return to the boundary value problem associated with an extremal arc E for the general problem of Mayer in the calculus of variations as introduced in §2, it is to be seen that if E is a normal non-singular extremal arc whose

end-values satisfy the end-conditions $\psi_\mu = 0$, and which satisfies condition I, then for the boundary value problem (9), (10) involving the $n + 2$ variables $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$ $(i = 1, \cdots, n)$ hypotheses ($H1$), ($H2$), ($H3$) and ($H5$) are satisfied. Whenever E is also assumed to satisfy the necessary condition of Clebsch for a minimum, it may be shown that there exists a constant λ_0 such that the expression

$$(34) \qquad g_2[\xi_1(x_1), \xi_2(x_2), \eta] + \lambda_0\{\xi_1^2(x_1) + \xi_2^2(x_2) + \int_{x_1}^{x_2} \eta_i(x)\, \eta_i(x)\, dx\}$$

is non-negative for all admissible sets $\xi_1(x)$, $\xi_2(x)$, $\eta_i(x)$. If in the derivation of the boundary value problem the quantity $g_2[\xi_1(x_1), \xi_2(x_2), \eta]$ is replaced by the expression (34), the modified differential system is equivalent to the original and may be reduced to it by a linear change of parameter. The modified boundary value problem satisfies the hypotheses ($H1$), ($H2$), ($H3$), ($H4$) and ($H5$), and in view of Theorem 6.1 has infinitely many characteristic numbers, all of which are seen to be positive. The original boundary value problem (9), (10) has, therefore, infinitely many characteristic numbers, of which only a finite number are negative; finally, Theorem 2.1 is a consequence of the minimizing property which characterizes the smallest characteristic number λ_1 of the system.

BIBLIOGRAPHY

I. Bliss, The problem of Mayer with variable end-points, Transactions of the American Mathematical Society, Vol. 19 (1918), pp. 305–314.

II. Cope, An analogue of Jacobi's condition for the problem of Mayer with variable end-points, Dissertation, The University of Chicago (1927).

III. Bliss, The problem of Lagrange in the calculus of variations, American Journal of Mathematics, Vol. 52 (1930), pp. 673–742.

IV. Morse and Myers, The problems of Lagrange and Mayer with variable end-points, Proceedings of the American Academy of Arts and Sciences, Vol. 66 (1931), pp. 235–253.

V. Morse, Sufficient conditions in the problem of Lagrange with variable end-conditions, American Journal of Mathematics, Vol. 53 (1931), pp. 517–546.

VI. Morse, Sufficient conditions in the problem of Lagrange with fixed end-points, Annals of Mathematics, Vol. 32 (1931), pp. 567–577.

VII. Bliss, The problem of Bolza in the calculus of variations, Annals of Mathematics, Vol. 33 (1932), pp. 261–274.

VIII. Bliss and Hestenes, Sufficient conditions for a problem of Mayer in the calculus of variations, Transactions of the American Mathematical Society, Vol. 35 (1933), pp. 305–326.

IX. Hestenes, Sufficient conditions for a problem of Mayer with variable end-points. [Dissertation, Chicago (1932)], Transactions of the American Mathematical Society, Vol. 35 (1933), pp. 479–490.

X. Reid, A boundary value problem associated with the calculus of variations, American Journal of Mathematics, Vol. 54 (1932), pp. 769–790.

The University of Chicago,
Chicago, Ill.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

SOME PROPERTIES OF EQUALITY AND IMPLICATION IN COMBINATORY LOGIC[1]

By H. B. Curry

(Received June 8, 1933)

This paper is an extension of, and is dependent on, my previous paper "Apparent Variables from the Standpoint of Combinatory Logic."[2] The extension consists in the fact that certain axioms for equality are introduced, and also certain properties of implication, hitherto introduced only as hypotheses in certain theorems, are explicitly formulated as axioms. The paper is concerned with the consequences of these new assumptions. The results are of two sorts, as follows:

1) A number of the simpler consequences of the new axioms are established in Theorems 1 to 10 inclusive. (These theorems are not necessary for the rest of the paper.) Among them is Theorem 7, by the aid of which it would be possible to restate some of the theorems in the paper referred to above without the requirement that certain formulas follow combinatorially.

2) In Theorems 11–14 a general result concerning the relation between certain kinds of theorems and their corresponding formulas is deduced. Hitherto we have often been concerned with the proof of a formula whose interpretation is some theorem previously established; the present Theorem 14 shows that it is possible, in a certain class of cases, to make this transition from theorem to formula in general. This class of cases occurs when we reason as follows:—"let $x_1, x_2, \cdots, x_n$ be such that $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_n$, hold, then .... $\mathfrak{X}$ holds"—the dots indicating an argument conducted according to the rules of combinatory logic, the x's treated formally. If such a theorem be proved, then Theorem 14 shows that it is possible to expand the proof into a rigorous[3] proof of a single formula expressing essentially the same fact; this formula is, for the case $m = 1$, the following

$$\vdash (x_1, x_2, \cdots, x_m)(\mathfrak{A} \supset \mathfrak{X}) .$$

If we call such a method of proof the *postulational method*, then the theorem under discussion justifies the postulational method as a method of establishing formulas in the same sense that my previous paper, referred to above, justified the use of apparent variables. Several of the theorems previously proved are,

---

[1] Presented to the American Math. Society, September 2, 1932. The paper is a product of the author's activities as National Research Fellow, 1931–32.

[2] See references to previous papers at end of this introduction.

[3] I.e., consisting of a series of inferences each of which is an application of a single one of the combinatory rules.

of course, special cases of this theorem; but the proof of the general result seems to require more axioms than was necessary for any of the special cases yet considered.

REFERENCES TO PREVIOUS PAPERS.

I. Grundlagen der kombinatorischen Logik, Kapitel I. *Am. Journ. Math.*, vol. 52, pp. 509–525 (1930).
II. Same, Kapitel II. *Am. Journ. Math.*, vol. 52, pp. 526–536, 789–834 (1930).
AT. Some additions to the theory of combinators. *Am. Journ. Math.*, vol. 54, pp. 551–558 (1932).
UQ. The universal quantifier in combinatory logic. *Annals of Math.*, Series 2, vol. 32, pp. 154–180 (1931).
AV. Apparent variables from the standpoint of combinatory logic. *Annals of Math.*, Series 2, vol. 34, pp. 381–404 (1931).

Axioms for Equality (including one previously introduced).

$$\text{Ax. } Q_0\text{:}\quad \vdash \Pi(W(CQ)) .$$

$$\text{Ax. } Q_1\text{:}\quad \vdash (x, y)(x = y \supset (x \supset y)) .$$

$$\text{Ax. } Q_2\text{:}\quad \vdash (x, y, z)(x = y \supset (zx \supset zy)) .$$

$$\text{Ax. } Q_3\text{:}\quad \vdash (x, y, z)((z)(xz = yz) \supset x = y) .$$

The first of these axioms was introduced under the name of Ax Q in I C 5; it was shown there to lead to the reflexive law of identity. The next two express principles related to the rules with corresponding symbols. As for the last one, it expresses a principle which has implicitly determined the course of many of my researches to date, indeed the combinatory axioms were chosen so that the fundamental theorem for combinators might hold. It is therefore not surprising that the restriction that certain formulas in AV follow combinatorially should fall when the axiom is assumed explicity.[4] The axiom is used below only in Theorems 6 and 7.

Axioms for Implication.

$$\text{Ax. (PB).}\quad \vdash (x, y, z)((x \supset y) \supset ((z \supset x) \supset (z \supset y))) .$$

$$\text{Ax. (PC).}\quad \vdash (x, y, z)((x \supset (y \supset z)) \supset (y \supset (x \supset z))) .$$

$$\text{Ax. (PW).}\quad \vdash (x, y)\quad ((x \supset (x \supset y)) \supset (x \supset y)) .$$

$$\text{Ax. (PK).}\quad \vdash (x, y)\quad (x \supset (y \supset x)) .$$

Theorem 1. $\vdash (x)(x = x)$.

Proof. See AV Th. 8.

Theorem 2. $\vdash (x, y)(x = y \supset y = x)$.

Proof.

[4] I strongly suspect that if, in addition to the axioms here postulated, axioms expressing the combinatory rules as formulas are assumed, then the combinatory axioms can be dispensed with entirely; but as yet I do not know whether this can be formally proved.

$$\vdash (x, y)(x = y \supset CQxx = CQxy) \qquad \left[\text{Subst.}\begin{pmatrix} CQx \\ z \end{pmatrix} \text{Ax. } Q_2\right].$$

$$\vdash (x, y)(CQxx = CQxy \supset (CQxx \supset CQxy)) \qquad [\text{Ax. } Q_1].$$

$$\therefore \vdash (x, y)(x = y \supset (CQxx \supset CQxy)) \qquad [\text{AV Th. } 14],$$

$$= \vdash (x, y)(x = y \supset (x = x \supset y = x)) \qquad [\text{Rule C; Def. of} =; \text{AV Th. } 4],$$

$$\supset \vdash (x, y)(x = x \supset (x = y \supset y = x))$$

$$\left[\text{Subst.}\begin{pmatrix} x = y, & x = x, & y = x \\ x & y & z \end{pmatrix} \text{Ax. (PC); AV Th. } 11\right].$$

$$\vdash (x, y)(x = x) \qquad [\text{Subst. Th. } 1].$$

$$\therefore \vdash (x, y)(x = y \supset y = x) \qquad [\text{AV Th. } 12] \text{ q.e.d.}$$

THEOREM 3. $\vdash (x, y, z)(x = y \supset (z = x \supset z = y))$.
Proof.

$$\vdash (x, y, z)(x = y \supset Qzx = Qzy) \qquad \left[\text{Subst.}\begin{pmatrix} Qz \\ z \end{pmatrix} \text{Ax. } Q_2\right],$$

$$= \vdash (x, y, z)(x = y \supset (z = x) = (z = y)) \qquad [\text{Def. of} =; \text{AV Th. } 4].$$

$$\vdash (x, y, z)((z = x) = (z = y) \supset (z = x \supset z = y)) \qquad [\text{Subst. Ax. } Q_1].$$

$$\vdash (x, y, z)(x = y \supset (z = x \supset z = y)) \qquad [\text{AV Th. } 14] \text{ q.e.d.}$$

THEOREM 4. $\vdash (x, y, z)(x = y \supset xz = yz)$.
Proof. By Subst. $\begin{pmatrix} CIz \\ z \end{pmatrix}$ Ax Q_2,

$$\vdash (x, y, z)(x = y \supset CIzx = CIzy)$$

$$= \vdash (x, y, z)(x = y \supset xz = yz) \qquad [\text{AV Th. } 4]. \quad \text{q.e.d.}$$

THEOREM 5. $\vdash (x, y, z_1, \cdots z_n)(x = y \supset x z_1 z_2 \cdots z_n = y z_1 z_2 \cdots z_n)$.
Proof. For $n = 1$, this is Theorem 4. Suppose this theorem proved for $n = k$, then it may be proved for $n = k + 1$ as follows: by this theorem for $n = k$

$$\vdash (x, y, z_2, z_3 \cdots z_{k+1})(x = y \supset xz_2 z_3 \cdots z_{k+1} = yz_2 z_3 \cdots z_{k+1}).$$

Hence, $\left[\text{Subst.}\begin{pmatrix} xz_1, & yz_1 \\ x\ , & y \end{pmatrix}\right]$

$$\vdash (x, y, z_1, z_2, \cdots z_{k+1})(xz_1 = yz_1 \supset xz_1 z_2 \cdots z_{k+1} = xz_1 z_2 \cdots z_{k+1}).$$

On the other hand,

$$\vdash (x, y, z_1, z_2 \cdots z_{k+1})(x = y \supset xz_1 = yz_1) \qquad [\text{Subst., Th. } 4].$$

The formula of this theorem for $n = k + 1$ follows from the last two formulas by AV Th. 14; whence the theorem follows by induction.

THEOREM 6. $\vdash (x, y)((z_1, z_2, \cdots z_n)(xz_1 z_2 \cdots z_n = yz_1 z_2 \cdots z_n) \supset x = y)$.

Proof. This follows from Ax. Q_3 in much the same way that the preceding theorem was derived from Theorem 4. In fact this theorem is true for $n = 1$ by Ax. Q_3. Let it be asumed for $n = k$, then it may be proved for $n = k + 1$ as follows: by this theorem for $n = k$,

$$\vdash (x, y)((z_2, z_3, \cdots, z_{k+1})(xz_2 z_3 \cdots z_{k+1} = yz_2 z_3 \cdots z_{k+1}) \supset x = y) .$$

Hence $\left[\text{Subst.} \begin{pmatrix} xz_1, & yz_1 \\ x\ , & y \end{pmatrix}\right]$

$$\vdash (x, y, z_1)((z_2, z_3 \cdots z_{k+1})(xz_1 z_2 \cdots z_{k+1} = yz_1 z_2 \cdots z_{k+1}) \supset xz_1 = yz_1) .$$

Therefore, [AV Ths. 11, 6, and 4]

$$\vdash (x, y)((z_1, z_2 \cdots z_{k+1})(xz_1 z_2 \cdots z_{k+1} = yz_1 z_2 \cdots z_{k+1}) \supset (z_1)(xz_1 = yz_1)) .$$

$$\vdash (x, y)((z_1)(xz_1 = yz_1) \supset x = y) \qquad [\text{Ax. } Q_3] .$$

The formula to be proved follows from the last two by AV Th. 14.

THEOREM 7. *If $\mathfrak{X}$ and $\mathfrak{Y}$ are combinations of constants and the variables $x_1, x_2, \cdots, x_n$, then a necessary and sufficient condition that*

(1) $$\vdash [x_1, x_2, \cdots, x_n]\, \mathfrak{X} = [x_1, x_2, \cdots, x_n]\, \mathfrak{Y} ,$$

is that

(2) $$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y}) .$$

Proof. Let $X \equiv [x_1, x_2, \cdots, x_n]\, \mathfrak{X}$,

and $Y \equiv [x_1, x_2, \cdots, x_n]\, \mathfrak{Y}$.

Then the condition (1) is

(3) $$\vdash X = Y ,$$

and by AV Ths. 2 and 4, (2) is equivalent to

(4) $$\vdash (x_1, x_2, \cdots, x_n)(Xx_1 x_2 \cdots x_n = Yx_1 x_2 \cdots x_n) .$$

That (4) follows from (3) is, however, a consequence of Theorem 5 and AV Th. 15; that (3) follows from (4) is a consequence of Theorem 6.

Remark. The necessity of the condition follows from Theorem 5, and so depends only on Ax. Q_2 and the axioms assumed prior to the present paper; on the other hand the sufficiency is a consequence of Ax. Q_3 and previous axioms.

THEOREM 8. *Ax. P_0 can be proved.*

Proof. $\vdash (x)(x = x \supset (x \supset x))$ $\qquad \left[\text{Subst.} \begin{pmatrix} x, x \\ x, y \end{pmatrix} \text{Ax. } Q_1\right]$.

$\therefore \vdash (x)(x \supset x)$. $\qquad$ [Th. 1; AV Th. 12]. q.e.d.

Ax. P_0 may also be derived from the axioms (PK) and (PC).

THEOREM 9. $\vdash (x, y)((z)(zx \supset zy) \supset x = y)$.

Proof. By AV Th. 17,

$$\vdash (x, y)((z)(zx \supset zy) \supset (Qxx \supset Qxy))$$

$$= \vdash (x, y)((z)(zx \supset zy) \supset (x = x \supset x = y))$$

$$\supset \vdash (x, y)(x = x \supset ((z)(zx \supset zy) \supset x = y)) \qquad [\text{Ax. PC; AV Th. } 12].$$

$$\therefore \vdash (x)(x = x \supset (y)((z)(zx \supset zy) \supset x = y)) \qquad [\text{AV Th. } 15],$$

from which the theorem follows by Theorem 1 and AV Theorems 12 and 6.

Remark. If we disregard the theory of types and similar complications then

$$[x, y]\ (z)(zx \supset zy)$$

is what equality is defined to be in the Principia Mathematica. This theorem and Ax. Q_2 shows that the relation just mentioned is formally equivalent to Q.

THEOREM 10. *If $\mathfrak{X}$ is a combination of constants and the variables $x, z_1, z_2, \cdots, z_n$, and $\mathfrak{Y}$ is an expression obtained from $\mathfrak{X}$ by substituting y for x in some or all—but not necessarily all—of its occurrences, then,*

$$\text{a)}\quad \vdash (x, y, z_1, z_2, \cdots, z_n)(x = y \supset \mathfrak{X} = \mathfrak{Y}).$$

$$\text{b)}\quad \vdash (x, y, z_1, z_2, \cdots, z_n)(x = y \supset (\mathfrak{X} \supset \mathfrak{Y})).$$

Proof. Let $\mathfrak{Z} \equiv [x, y]\,\mathfrak{Y}$.

Then $\vdash^{c}_{x,y,z} \mathfrak{Y} = \mathfrak{Z}xy \qquad [\text{AV Th. } 2].$

$$\vdash^{c}_{x\,z} \quad \mathfrak{X} = \mathfrak{Z}xx \qquad [\text{AV Th. } 5].$$

$$\vdash (x, y, z_1, \cdots, z_n)(x = y \supset \mathfrak{Z}xx = \mathfrak{Z}xy) \qquad \left[\text{Subst.} \begin{pmatrix} \mathfrak{Z}x \\ z \end{pmatrix} \text{Ax. } Q_2\right],$$

$$= \vdash (x, y, z_1, \cdots, z_n)(x = y \supset \mathfrak{X} = \mathfrak{Y}) \qquad [\text{AV Th. } 4].$$

which proves a).

Again, $\left[\text{Subst.} \begin{pmatrix} \mathfrak{X}, \mathfrak{Y} \\ x, y \end{pmatrix} \text{Ax. } Q_1\right]$

$$\vdash (x, y, z_1, z_2, \cdots, z_n)(\mathfrak{X} = \mathfrak{Y} \supset (\mathfrak{X} \supset \mathfrak{Y}))$$

$$\supset \vdash (x, y, z_1, z_2, \cdots, z_n)((x = y \supset \mathfrak{X} = \mathfrak{Y}) \supset (x = y \supset (\mathfrak{X} \supset \mathfrak{Y})))$$

$$\left[\text{Subst.} \begin{pmatrix} \mathfrak{X} = \mathfrak{Y}, & \mathfrak{X} \supset \mathfrak{Y}, & x = y \\ x\quad, & y\quad, & z \end{pmatrix} \text{Ax. (PB); AV Th. } 12\right].$$

From this we have b) by a) and AV Th. 12.

THEOREM 11. *If $\mathfrak{X}$ and $\mathfrak{Y}$ are combinations of constants and the variables $x_1, x_2, \cdots x_n$, such that*

$$\text{i)}\ \vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{X},$$

$$\text{ii)}\ \vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y});$$

then

$$\vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{Y}.$$

Proof.

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y} \supset (\mathfrak{X} \supset \mathfrak{Y})) \qquad \left[\text{Subst.} \begin{pmatrix} \mathfrak{X}, \mathfrak{Y} \\ x, y \end{pmatrix} \text{Ax. } Q_1\right].$$

$$\therefore \vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} \supset \mathfrak{Y}) \qquad [\text{Hp. ii; AV Th. 12}].$$

$$\therefore \vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{Y}. \qquad [\text{Hp. 1; Av Th. 12}]. \quad \text{q.e.d.}$$

Remark. If Ax. Q_3 be assumed, the theorem follows immediately from Theorem 7, AV Def. 4, and Rule Q_1.

CONVENTION 1. *The term $\mathfrak{L}$ system shall denote any formal theory such that:*

(i) *it has the same*[5] *non-formal primitive ideas as those stated for combinatory logic in* I C 2;

(ii) *the formal primitive ideas include Q, B, C, W, K, P, Π,*

(iii) *the rules consist of the rules E, Q_1, Q_2, B, C, W, K, P, Π, of I C* 6, *and no others (except possibly such as can be derived from these as theorems);*

(iv) *the axioms include all those postulated for combinatory logic in I, UQ, AV, and the present paper, except that Q_3 is not required to be included.*

The definitions and conventions which have been set down and the theorems which have been proved for combinatory logic will be extended, where they apply, to an arbitrary $\mathfrak{L}$ system.

CONVENTION 2. *If $\mathfrak{S}$ is an $\mathfrak{L}$ system, then*

$$|\mathfrak{S}\vdash X$$

shall denote that X is an assertion in $\mathfrak{S}$. Similarly the symbols

$$|\mathfrak{S}\vdash_x, \quad = |\mathfrak{S}\vdash, \quad \supset |\mathfrak{S}\vdash, \quad \text{etc.}$$

shall have meanings analogous to those of $\vdash_x$, etc.

The above notation will be used when two or more systems are simultaneously

[5] This sameness is an identity in names. If the reader prefers he may define an $\mathfrak{L}$ system as one which can be put into a partial isomorphism with the portion of combinatory logic described. But in that case we can give corresponding elements the same names, and we then have the same definition as in the text. In this paper and also in some others, I have not always upheld the distinction between symbols and the entities they represent; it is doubtful whether, in an abstract theory, there is any such distinction.

under discussion. Where not ambiguous the simple sign $\vdash$ etc., will be used, referring to the single system being considered.

THEOREM 12. *If $\mathfrak{S}$ is an $\mathfrak{L}$ system (with the omission perhaps of the axioms PC, PW, and PK) not containing the symbols $x_1, x_2, \cdots, x_n$ as representing primitive entities, and $\mathfrak{S}^*$ is the system obtained by adjoining $x_1, x_2, \cdots, x_n$ to the list of formal primitive ideas of $\mathfrak{S}$; then*

a) *the entities of $\mathfrak{S}^*$ are the arbitrary combinations of entities of $\mathfrak{S}$ and $x_1, x_2, \cdots, x_n$;*

b) *If $\mathfrak{X}$ is an entity of $\mathfrak{S}^*$ then a necessary and sufficient condition that*

$$|\mathfrak{S}^* \vdash \mathfrak{X} \tag{1}$$

is that

$$|\mathfrak{S} \vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{X}\,. \tag{2}$$

Proof. a) follows at once; for the Rule E, which is essentially a definition of entity in $\mathfrak{S}^*$, is the same as the definition of the stated type of combination (II A, Festsetzung 1).

As for b), we observe first that (2) is sufficient; for if $(x_1, x_2, \cdots, x_n)\mathfrak{X}$ is an assertion in $\mathfrak{S}$, it is *ipso facto* an assertion in $\mathfrak{S}^*$; and from the formula

$$|\mathfrak{S}^* \vdash (x_1, x_2, \cdots, x_n)\mathfrak{X}$$

we obtain (1) by Subst. in $\mathfrak{S}^*$ ($x_1, \cdots x_n$ are entities in $\mathfrak{S}^*$).

Next let X be an assertion in $\mathfrak{S}$, i.e., such that

$$|\mathfrak{S} \vdash X\,. \tag{3}$$

Let

$$V_n \equiv K_n \cdot K_{n-1} \cdot K_{n-2} \cdots K_2 \cdot K_1\,.$$

Then

$$\vdash_x^c V_n X x_1 x_2 \cdots x_n = X$$

and hence [AV Th. 4, Cor. 1]

$$\vdash^c (x_1, x_2, \cdots, x_n)\, X = \Pi_n(V_n X)\,.$$

But V_n is a regular combinator of order n and degree 0. Hence by (3) and UQ 7, Th. 6,

$$\begin{aligned} &|\mathfrak{S} \vdash \Pi_n(V_n X) \\ = &|\mathfrak{S} \vdash (x_1, x_2, \cdots, x_n)\, X\,. \end{aligned}$$

Thus (2) is true whenever $\mathfrak{X}$ is itself an assertion in $\mathfrak{S}$, in particular—since the primitive assertions of $\mathfrak{S}$ and $\mathfrak{S}^*$ are the same—if the formula (1) represents an axiom of $\mathfrak{S}^*$.

Now if $\mathfrak{X}$ is an arbitrary non-primitive assertion of $\mathfrak{S}^*$, then there will exist a

proof of the formula (1), which proof will consist of a series of steps, each of which is a single application of some rule of $\mathfrak{S}^*$ whereby, starting from the axioms of $\mathfrak{S}^*$, we arrive at (1). Suppose now we replace all the formulas of type (1) appearing in these rules by the corresponding formulas of type (2); then we shall have a set of statements describing certain conceivable inferences in $\mathfrak{S}$. If these statements can be shown to be true theorems concerning $\mathfrak{S}$—i.e., if it be shown how, in any particular case, the inference in question can be accomplished by a series of elementary inferences, each an application of one of the rules of $\mathfrak{S}$—then the proof of the formula (1) can be transformed step by step into a proof of (2). For the analogues of type (2) of the initial formulas of this proof, are true by the preceding paragraph; and if the same be true for all the formulas up to any stage in the original proof, then it will be true for the next stage, since the rule by which the inference was originally made becomes a valid inference from previously proved formulas of type (2) to the analogue of the next formula in the original proof. Thus the proof of the necessity of the condition in b) will be complete as soon as the above supposition concerning the rules of $\mathfrak{S}^*$ is substantiated.

This supposition will be proved seriatim for the various rules of $\mathfrak{S}^*$ below. In this discussion the analogues, according to the above method, of the rules of $\mathfrak{S}^*$ will be distinguished by a star from the corresponding rules of $\mathfrak{S}$. The statement of the transformed rule is given first, then its proof. All symbols $\vdash$ refer to the system $\mathfrak{S}$.

Rule Q_1^*. If $\mathfrak{X}$ and $\mathfrak{Y}$ are combinations such that $\vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{X}$ *and*

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y}),$$

then

$$\vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{Y}.$$

This was proved in Theorem 11.

Rule Q_2^*. If $\mathfrak{X}$, $\mathfrak{Y}$, and $\mathfrak{Z}$ are combinations and if

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y}),$$

then

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{Z}\mathfrak{X} = \mathfrak{Z}\mathfrak{Y}).$$

For by Subst. $\begin{pmatrix} \mathfrak{X}, \mathfrak{Y}, \mathfrak{Z} \\ x, y, z \end{pmatrix}$ Ax. Q_2,

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} = \mathfrak{Y} \supset \mathfrak{Z}\mathfrak{X} = \mathfrak{Z}\mathfrak{Y}),$$

whence the rule follows by AV Th. 12.

Rules B^*, C^*, W^*, K^*. If $\mathfrak{X}$, $\mathfrak{Y}$, $\mathfrak{Z}$ are combinations, then

$$\vdash (x_1, x_2, \cdots, x_n)(\mathrm{B}\mathfrak{X}\mathfrak{Y}\mathfrak{Z} = \mathfrak{X}(\mathfrak{Y}\mathfrak{Z})).$$

$$\vdash (x_1, x_2, \cdots, x_n)(\mathrm{C}\mathfrak{X}\mathfrak{Y}\mathfrak{Z} = \mathfrak{X}\mathfrak{Z}\mathfrak{Y}).$$

$$\vdash (x_1, x_2, \cdots, x_n)(\mathrm{W}\mathfrak{X}\mathfrak{Y} = \mathfrak{X}\mathfrak{Y}\mathfrak{Y}).$$

$$\vdash (x_1, x_2, \cdots, x_n)(\mathrm{K}\mathfrak{X}\mathfrak{Y} = \mathfrak{X}).$$

These rules follow by AV Th. 8.

*Rule P**. *If* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *are combinations, such that* $\vdash (x_1, x_2, \cdots, x_n)\,\mathfrak{X}$ *and*

$$\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{X} \supset \mathfrak{Y}),$$

then

$$\vdash (x_1, x_2, \cdots, x_n)\mathfrak{Y}.$$

This is AV Th. 12.

Rule Π*. *If* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *are combinations, and if*

$$\vdash (x_1, x_2, \cdots, x_n)\, \Pi\mathfrak{X};$$

then

$$\vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{X}\mathfrak{Y}.$$

For by Subst. $\begin{pmatrix}\mathfrak{X}, \mathfrak{Y}\\ x, y\end{pmatrix}$ Ax. Π_0.

$$\vdash (x_1, x_2, \cdots, x_n)(\Pi\mathfrak{X} \supset \mathfrak{X}\mathfrak{Y}),$$

whence the rule follows by AV Th. 12.

COROLLARY 1. *If* $\mathfrak{S}$ *is as in the theorem, then a necessary and sufficient condition that*

(4) $$|\mathfrak{S} \vdash_x \mathfrak{X},$$

is that $|\mathfrak{S} \vdash (x_1, x_2, \cdots, x_n)\mathfrak{X}$.

Proof. Follows from the theorem, since by definition (4) is only another way of writing (1).

COROLLARY 2. *If* $\mathfrak{S}$ *is as in the theorem, except that Axioms* Q_1 *and* Q_2 *need not be asserted in* $\mathfrak{S}$, *and* $\mathfrak{X}$ *is an expression such that*

(4) $$|\mathfrak{S} \vdash_x \mathfrak{X}$$

follows with the restriction that formulas of the type

$$|\mathfrak{S} \vdash_x \mathfrak{U} = \mathfrak{V}$$

shall be used in the proof only when they follow combinatorially, then

(5) $$|\mathfrak{S} \vdash_x (x_1, x_2, \cdots, x_n)\, \mathfrak{X}.$$

Proof. It is necessary only to revise the proofs of Rule Q_1^* and Q_2^* above, since these are the only places where the excepted axioms Q_1 and Q_2 were used.

Suppose then we have a use of Rule Q_1^* in the original proof of (4). By the hypotheses of this corollary this must be in form as follows: "If $\vdash_x \mathfrak{U}$ and $\vdash_x^c (\mathfrak{U} = \mathfrak{V})$, then $\vdash_x \mathfrak{V}$." Then if we prove $\vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{U}$, we have $\vdash (x_1, \cdots, x_n)\, \mathfrak{V}$ by AV Th. 4.

Again suppose there is a use of Rule Q_2 in the proof of (4). This must be as follows: If $\vdash_x^c (\mathfrak{U} = \mathfrak{V})$ then $\vdash_x^c (\mathfrak{Z}\mathfrak{U} = \mathfrak{Z}\mathfrak{V})$." In that case the corresponding

conclusion in the proof of (5), viz., $\vdash (x_1, x_2, \cdots, x_n)(\mathfrak{Z}\mathfrak{U} = \mathfrak{Z}\mathfrak{V})$, follows from AV Th. 8.

If we make these changes in the proof above it will still follow that a proof of (4) can be transformed step by step into a proof of (5).

THEOREM 13. *If $\mathfrak{S}$ is an $\mathfrak{L}$ system, A an entity of $\mathfrak{S}$, and $\mathfrak{S}'$ the system obtained by adjoining A to the list of primitive assertions of $\mathfrak{S}$; then*

a) *the entities of $\mathfrak{S}$ are the same as those of $\mathfrak{S}'$, and*

b) *for any entity X a necessary and sufficient condition that*

$$(1) \qquad | \mathfrak{S}' \vdash X$$

is that

$$(2) \qquad | \mathfrak{S} \vdash A \supset X .$$

Proof. The statement a) I take to be self-evident.

b) The condition is sufficient; for if (2) holds, so does

$$| \mathfrak{S}' \vdash A \supset X ,$$

from which we obtain (1) by Hp. and Rule P (in $\mathfrak{S}'$).

If X is a primitive assertion (axiom) in $\mathfrak{S}'$ then it is either A itself, in which case (2) follows by Theorem 8, or else it is an axiom of $\mathfrak{S}$, in which case (2) follows by Subst. $\begin{pmatrix} X, A \\ x, y \end{pmatrix}$ Ax. (PK). Thus (2) holds whenever (1) is an axiom of $\mathfrak{S}'$.

As in the proof of Theorem 12 the necessity of the condition (2) will be established if it be shown that the rules of $\mathfrak{S}'$ become valid theorems concerning $\mathfrak{S}$ when assertion in $\mathfrak{S}'$ is interpreted in the sense (2). This will be done as follows (All formulas refer to $\mathfrak{S}$, cf. proof of Theorem 12).

Rule Q_1'. "If X and Y are entities and

$$\vdash A \supset X ,$$
$$\vdash A \supset (X = Y) ,$$

then

$$\vdash A \supset Y."$$

For

$$\vdash X = Y \supset (X \supset Y) \qquad [\text{Ax. } Q_1] ,$$
$$\supset \vdash (A \supset X = Y) \supset (A \supset (X \supset Y)) \qquad [\text{Ax. (PB)}] .$$
$$\therefore \vdash A \supset (X \supset Y) \qquad [\text{Hp.}] .$$

The rule in question has been reduced to that discussed below under Rule P′.

Rule Q_2'. "If X, Y and Z are entities, and

$$\vdash A \supset (X = Y) ,$$

then

$$\vdash A \supset (ZX = ZY)\text{ .''}$$

For $\vdash X = Y \supset ZX = ZY$ [Ax. Q_2]

$\supset \vdash (A \supset (X = Y)) \supset (A \supset (ZX = ZY))$ [Ax. (PB)],

whence the rule follows by Rule P.

Rules B′, C′, W′, K′. "*If X, Y, Z are any entities, then*

$$\vdash A \supset BXYZ = X(YZ),$$

$$\vdash A \supset CXYZ = XZY,$$

$$\vdash A \supset WXY = XYY,$$

$$\vdash A \supset KXY = X\text{ .''}$$

For the formulas follow, by Ax. (PK), from those established by Rules B, C, W, K.

Rule P′. "*If X and Y are entities such that*

$$\vdash A \supset X \quad \textit{and} \quad \vdash A \supset (X \supset Y)\,;$$

then

$$\vdash A \supset Y\text{ .''}$$

For $\vdash (X \supset Y) \supset ((A \supset X) \supset (A \supset Y))$ [Ax. (PB)],

$\supset \vdash (A \supset X) \supset ((X \supset Y) \supset (A \supset Y))$ [Ax. (PC)].

$\therefore \vdash (X \supset Y) \supset (A \supset Y)$ [Hp., Rule P],

$\supset \vdash (A \supset (X \supset Y)) \supset (A \supset (A \supset Y))$ [Ax. (PB)].

$\therefore \vdash A \supset (A \supset Y)$ [Hp.; Rule P],

$\supset \vdash A \supset Y$ [Ax. (PW)].

Rule Π′. "*If X and Y are entities and*

$$\vdash A \supset \Pi X;$$

then

$$\vdash A \supset X Y\text{ .''}$$

For $\vdash \Pi X \supset X Y$ [Ax. Π_0].

$\supset \vdash (A \supset \Pi X) \supset (A \supset X Y)$ [Ax. (PB)],

so that the rule in question follows by Rule P.

THEOREM 14. *If i) $\mathfrak{S}$ is an $\mathfrak{L}$ system not containing the variables $x_1, x_2, \cdots, x_n$, ii) $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_m$ are combinations of entities in $\mathfrak{S}$ and the variables $x_1, x_2, \cdots, x_n$, iii) $\mathfrak{T}$*

is the system formed by adjoining to $\mathfrak{S}$ *the variables* $x_1, x_2, \cdots, x_n$ *as primitive ideas and the* $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_m$ *as primitive assertions; then a necessary and sufficient condition that*

$$\mid \mathfrak{T} \vdash \mathfrak{X}\,, \tag{1}$$

is that

$$\mid \mathfrak{S} \vdash (x_1, x_2, \cdots, x_n)\, \mathfrak{Z}_m\,, \tag{2}$$

where $\mathfrak{Z}_1, \mathfrak{Z}_2, \cdots, \mathfrak{Z}_m$ *are defined by recursion as follows:*

$$\mathfrak{Z}_0 \equiv \mathfrak{X}\,, \qquad \mathfrak{Z}_{k+1} \equiv \mathfrak{A}_{m-k} \supset \mathfrak{Z}_k\,, \qquad k = 0, 1, 2, \cdots, m-1\,. \tag{3}$$

Proof. Let $\mathfrak{S}_k^*$ $(k = 0, 1, 2, \cdots, m)$ be the system formed by adjoining to $\mathfrak{S}$ the $x_1, x_2, \cdots, x_n$ as primitive ideas, and $\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_k$ as primitive assertions. Then $\mathfrak{S}_m^*$ is the same as $\mathfrak{T}$, and $\mathfrak{S}_{k+1}^*$ is formed by adjoining $\mathfrak{A}_{k+1}$ as primitive assertion to $\mathfrak{S}_k^*$. The formula (1) is equivalent to

$$\mid \mathfrak{S}_{m-k}^* \vdash \mathfrak{Z}_k \tag{4}$$

for $k = 0$, while (2) is, by Theorem 12, equivalent to (4) for $k = m$.

On the other hand, since $\mathfrak{S}_{m-k}^*$ is formed from $\mathfrak{S}_{m-k-1}^*$ by the adjunction as primitive assertion of $\mathfrak{A}_{m-k}$, it follows by Theorem 13 that (4) is, for any fixed k, equivalent to

$$\begin{aligned} &\mid \mathfrak{S}_{m-k-1}^* \vdash \mathfrak{A}_{m-k} \supset \mathfrak{Z}_k \\ = &\mid \mathfrak{S}_{m-k-1}^* \vdash \mathfrak{Z}_{k+1} \qquad \text{[by (3)]}\,, \end{aligned}$$

which is the same formula (4) with k increased one unit. Thus (4) for a given k, is equivalent to the same formula for $k + 1$; hence the formulas (4) for $k = 0$ and $k = m$ respectively are equivalent, q.e.d.

UNIVERSITY OF CHICAGO.

ANNALS OF MATHEMATICS
Vol. 35, No. 4, October, 1934

THE TOPOLOGY OF TRANSFORMATION-SETS

BY GARRETT BIRKHOFF[1]

(Received November 25, 1933)

1. **Introduction.** Tietze[2] observed that the (1,1) continuous transformations of any space into itself constitute a group, of which the deformations are a normal subgroup. Baer[3] has shown that this group can be assigned continuity relations; as a matter of fact, this can be done in many ways, of which we shall discuss three in the text.

Schreier[4] pointed out that in a continuous group, there should be harmony between group operations and limit operations; Kneser[5] has developed this notion in outline for Tietze's groups. Finally, Fréchet and later writers[6] have shown that the full meaning of continuity is had only by spaces homeomorphic with subsets of Hilbert space.

We shall try to unite these tendencies of thought by showing how the (1,1) continuous transformations of a space $\mathfrak{S}$ into itself can be regarded as a group (i) continuous in the sense of Schreier (ii) topologically equivalent to a subset of Hilbert space.

2. **Outline.** In carrying out the program projected in section one, it has been found essential to proceed with great care. For at every step we have a variety of definitions to choose from, and it is only after a detailed examination that we can judge of their relative fruitfulness.

Accordingly, in §§3–9 we lay the groundwork for the subsequent treatment by putting the properties of finite sets of transformations in the most general possible form. In §§10–15 we turn our attention to infinite sequences of transformations, to see the extent to which various convergence definitions yield the classical convergence properties; the answer to this is seen to hinge on the convergence properties of the spaces transformed.

In §§16–20 the adequacy of three convergence definitions from the point of

[1] Society of Fellows, Harvard University.

[2] "Über die topologische Invarianten mehrdimensionaler Mannigfaltigkeiten," Monatsh. f. Math. u. Phys., **19** (1908), esp. pp. 88–93.

[3] "Beziehungen zwischen den Grundbegriffen der Topologie," Heidel. Sitz. 1929, (13). Cf. also Kneser's article.

[4] "Abstrakte kontinuierliche Gruppen," Hamb. Abh. **4** (1926), 15–32.

[5] "Die Deformationssätze der einfach zusammenhängenden Flächen," Math. Zeits. **25** (1926), 362–79.

[6] Notably Hausdorff, Urysohn, and Tychonoff. For the author's point of view, cf. "Axiomatic definitions of perfectly separable metric spaces," Bull. Am. Math. Soc., **39** (1933), 601–7.

view of Schreier and Kneser is discussed; incidentally we get considerable information about their equivalence. But it is in §§21–23 that we reach the core of the paper—namely, the fact there are important transformation-sets to which we can apply simultaneously the theory of abstract continuous groups, and the theory of topology.

3. **Transformations.** Let S be a class of objects s_k, and T a class of objects t_k. By a "transformation of S into T" is meant a rule σ which assigns to each couple (s_i, t_j) one of the following locutions

(i) σ transforms s_i into t_j.

(ii) σ does not transform s_i into t_j.

By the "inverse" σ^{-1} of σ, is meant that transformation of T into S which transforms a given t_i of T into an arbitrary s_j of S if and only if σ transforms s_j into t_i. This relation is *reciprocal*; that is, $(\sigma^{-1})^{-1} = \sigma$.

σ is called $(1 \to 1)$ or "one-valued" if and only if, given s_i, there is just one t_j such that σ assigns the locution (i) to (s_i, t_j). We write this relation in symbols as $t_j = \sigma(s_i)$, or $\sigma(s_i) = t_j$.

σ is called $(1,1)$ or "one-to-one" if and only if σ and σ^{-1} are both one-valued. If there exists a $(1,1)$ transformation of S into T, then the cardinal number of S equals the cardinal number of T. Because of the reciprocal nature of σ and σ^{-1}, if σ is $(1,1)$, then so is σ^{-1}.

4. **Products.** Let σ be any transformation of the class S into the class T, and τ any transformation of T into the class U. By the "product" $\sigma\tau$ of σ and τ, is meant that transformation of S into U which transforms s_i of S into u_j of U if and only if t_k of T exists, such that σ transforms s_i into t_k, while τ transforms t_k into u_j.

The reader will find it easy to construct proofs, in terms of the above definitions, of the following facts:

THEOREM 1: *$\tau^{-1}\sigma^{-1}$ is the inverse of $\sigma\tau$.*

THEOREM 2: *If σ and τ are both $(1 \to 1)$, then so is $\sigma\tau$.*

COROLLARY: *If σ and τ are both $(1,1)$, then so is $\sigma\tau$.*

THEOREM 3: $(\rho\sigma)\tau = \rho(\sigma\tau)$.

5. **The Identity.** Let S be any class. By 1 is meant that $(1,1)$ transformation of S into itself which transforms each s_i of S into itself and into itself alone. Obviously

THEOREM 4: *The identity 1 is its own inverse. And if σ is any transformation of S into T, then $1\sigma = \sigma 1 = \sigma$.*

If σ is a transformation of S into T, then $\sigma\sigma^{-1} = 1$ if and only if σ transforms each s_i of S into *some* t_j of T, and no two distinct s_i of S into the same t_j of T. While $\sigma^{-1}\sigma = 1$ if and only if each t_j of T is a transform of *some* s_i of S, and no S_i of S is carried into two distinct t_j of T. Repairing the four conditions just obtained, we see

THEOREM 5: $\sigma\sigma^{-1} = \sigma^{-1}\sigma = 1$ *if and only if σ is* $(1,1)$.

6. **Groups and Groupoids.** By a "continuous group" is meant[7] any system $\mathfrak{G}$ whose elements satisfy the postulates

G1: Every ordered pair (a, b) of elements of $\mathfrak{G}$ has a *product* ab in $\mathfrak{G}$.

G2: $(ab)c = a(bc)$ for any a, b, c in $\mathfrak{G}$.

G3: $\mathfrak{G}$ contains an element 1 satisfying $a1 = 1a = a$ for every a in $\mathfrak{G}$.

G4: To every a of $\mathfrak{G}$ corresponds an *inverse* a^{-1} satisfying $aa^{-1} = a^{-1}a = 1$.

L1: Concerning any element a and any enumerable sequence of elements a_n of $\mathfrak{G}$ we can say either (i) $\{a_n\}$ converges to a or (ii) $\{a_n\}$ does not converge to a. (In symbols, $a_n \to a$ or $a_n \nrightarrow a$).

LG1: If $a_n \to a$ and $b_n \to b$, then $a_n b_n \to ab$.

LG2: If $a_n \to a$, then $a_n^{-1} \to a^{-1}$.

A system $\mathfrak{G}$ is called a "group" provided only it satisfies G1–G4; a "continuous groupoid" provided it satisfies G1–G3, L1, and LG1; a "groupoid" provided it satisfies G1–G3. Definitions of subgroupoid and subgroup can be easily supplied by the reader.

We can summarize a number of the facts already stated in the following

THEOREM 6: *The transformations of any class into itself are a groupoid, of which the* $(1 \to 1)$ *transformations are a subgroupoid, and the* $(1, 1)$ *transformations a sub-subgroupoid and also a group.*

7. **Abstract Space.** Let Γ: $(\mathfrak{C}, K)$ be any abstract space, that is,[8] a class $\mathfrak{C}$ of points, and an operation K which assigns to each set S of points of $\mathfrak{C}$ a "derived" set $K(S)$.

Let $\mathfrak{S}$ be any set of points of Γ. By the "space" of $\mathfrak{S}$, is meant that abstract space Σ: $(\mathfrak{S}, L)$ in which L assigns to each set S of points of $\mathfrak{S}$, the "derived" set $L(S)$ consisting of those points of $K(S)$ in $\mathfrak{S}$.

Disregarding for the moment the operation K, it is obvious that we can speak of the transformations of Γ into itself, of Γ_1 into Γ_2, etc. But this is fruitless unless we tie up these transformations with the operation of derivation.

In doing this, it is convenient[9] to restrict ourselves to $(1 \to 1)$ transformations. As we have seen, a $(1 \to 1)$ transformation of the abstract space Γ_1: $(\mathfrak{C}_1, K_1)$ into the abstract space Γ_2: $(\mathfrak{C}_2, K_2)$ is a rule γ assigning to each point x in $\mathfrak{C}_1$ an "image" $\gamma(x)$ in $\mathfrak{C}_2$. It follows immediately that γ assigns to every set S of points of $\mathfrak{C}_1$, an "image" $\gamma(S)$ in $\mathfrak{C}_2$, namely, the set of the images of the points of S.

γ can also be regarded as a transformation of Γ_1 into the space of $\gamma(\mathfrak{C}_1)$, or into the space of any subset of $\mathfrak{C}_2$ which contains $\gamma(\mathfrak{C}_1)$.

The following result is obvious

THEOREM 7: $\sigma\tau(S) = \tau(\sigma(S))$.

---

[7] Cf. O. Schreier, op. cit.; also an article on "Hausdorff Groupoids" by the author, *Annals of Mathematics*, **35**, pp. 351–360.

[8] M. Fréchet, "Les espaces abstraits," Paris, 1926, p. 167. K is evidently a $(1 \to 1)$ transformation of the class of the S into itself.

[9] But by no means necessary; all the definitions given below can be so worded as to apply to the most general transformation between two abstract spaces.

8. **Continuous** $(1 \to 1)$ **Transformations.** A $(1 \to 1)$ transformation γ of Γ_1 into Γ_2 is said to be *continuous*[10] if and only if $\gamma(K_1(S)) \subset K_2(\gamma(S))$ for any set S of points of $\mathfrak{C}_1$.

THEOREM 8: *If Σ is the space defined by the set $\mathfrak{S}$ of Γ_2, and γ is any continuous $(1 \to 1)$ transformation of Γ_1 into Σ, then γ is a continuous $(1 \to 1)$ transformation of Γ_1 into Γ_2.*

Because, formally, $\gamma(K_1(S)) \subset L(\gamma(S)) \subset K_2(\gamma(S))$ for any subset S of $\mathfrak{C}_1$.

THEOREM 9: *The product of two continuous $(1 \to 1)$ transformations is itself a continuous $(1 \to 1)$ transformation.*

That it is $(1 \to 1)$ follows from Theorem 2. To show that it is continuous, suppose γ_1: $\Gamma_1 \to \Gamma_2$ and γ_2: $\Gamma_2 \to \Gamma_3$ to be the given transformations, and let S be any point-set in $\mathfrak{C}_1$. Then

$$\begin{aligned} K_3(\gamma_1\gamma_2(S)) &= K_3(\gamma_2(\gamma_1(S))) \supset \gamma_2(K_2(\gamma_1(S))) \\ &\supset \gamma_2(\gamma_1(K_1(S))) = \gamma_1\gamma_2(K_1(S)) \,. \end{aligned}$$

COROLLARY: *The continuous $(1 \to 1)$ transformations of any abstract space into itself are a groupoid.*

In the case that the K_i are defined by convergent sequences,[11] the criterion for continuity given above amounts to requiring that $\{x_k\} \to x$ imply $\{\tau(x_k)\} \to \tau(x)$.

9. **Bicontinuous** $(1,1)$ **Transformations.** A $(1,1)$ transformation γ is said to be *bicontinuous* if and only if γ and γ^{-1} are both[12] continuous in the sense of §8. Suppose then γ, a bicontinuous $(1,1)$ transformation of Γ_1 into Γ_2. Irrespective of $S \subset \mathfrak{C}_1$, we will have

$$\begin{aligned} \gamma(K_1(S)) &\overset{\subset}{\supset} K_2(\gamma(S)) \quad \text{and} \\ \gamma(K_1(S)) &= \gamma(K_1(\gamma^{-1}(\gamma(S)))) \\ &\supset \gamma(\gamma^{-1}(K_2(\gamma(S)))) = K_2(\gamma(S)) \,. \end{aligned}$$

Whence $\gamma(K_1(S)) = K_2(\gamma(S))$. From this we can further infer

$$\begin{aligned} \gamma^{-1}(K_2(S)) &= \gamma^{-1}(K_2(\gamma(\gamma^{-1}(S)))) \\ &= \gamma^{-1}(\gamma(K_1(\gamma^{-1}(S)))) = K_1(\gamma^{-1}(S)) \,. \end{aligned}$$

Moreover these conditions are obviously sufficient to ensure that γ be bicontinuous, giving us

THEOREM 10: *A necessary and sufficient condition that γ be bicontinuous is that $\gamma(K_1(S)) = K_2(\gamma(S))$ irrespective of S.*

---

[10] Cf., e.g., "Les espaces abstraits," p. 177. We are using throughout the notation Γ_i: $(\mathfrak{C}_i, K_i)$.

[11] That is, that the Γ_i satisfy (L1) of §6, and $K(S) \supset p$ if and only if S contains a sequence converging to p.

[12] Continuity does not imply bicontinuity, even in the most regular spaces, as we could easily show by examples.

In other words, a necessary and sufficient condition that a (1,1) transformation γ of an abstract space into itself be bicontinuous is that γ commute with the operation of derivation.

Because of the reciprocal nature of γ and γ^{-1}, it is at once evident that if γ is bicontinuous, then so is γ^{-1}. Moreover by Theorems 7 and 9, if γ and γ' are bicontinuous, then so is $\gamma\gamma'$. Since finally $\gamma\gamma^{-1} = \gamma^{-1}\gamma = 1$, and $1\gamma = \gamma = \gamma 1$, we have

THEOREM 11: *The* (1,1) *bicontinuous transformations of any abstract space into itself constitute a subgroup of the group of all* (1,1) *transformations of the space into itself.*

A bicontinuous (1,1) transformation amounts by Theorem 10 to a homeomorphic correspondence, which is perhaps the most basic concept in abstract topology. The advantage of our definition is that it correlates the various types of transformations worked with in a single pattern.

Theorem 11 was, as has been remarked, asserted by Tietze, but its explicit content has to the author's knowledge been unexplained hitherto, and least of all in relation to general abstract spaces.

In terms of convergence, a transformation τ is bicontinuous if and only if $p_k \to p$ implies $\tau(p_k) \to \tau(p)$, and conversely.

10. **Types of Convergence Space.** By a *convergence space*,[13] will be meant any set of points satisfying postulate L1 of §6. A convergence space will be called an *S-space* if and only if the following four single limit properties hold,

SL1: If $x_k = x$, then $x_k \to x$.
SL2: If $x_k \to x$ and $x_k \to y$, then $x = y$.
SL3: If $k(i) \to \infty$ $[i = 1, 2, 3, \cdots]$ and $x_k \to x$, then $x_{k(i)} \to x$.
SL4: If to any $k(i) \to \infty$ corresponds an $n(i) \to \infty$ satisfying $x_{k(n(i))} \to x$, then $x_k \to x$.

An S-space will be called a *D-space* if and only if the following two double limit properties hold,

DL1: If $x_j^i \to x_i$ for every positive integer i, and $x_i \to x$, then $N(i)$ exists so large that $j(i) > N(i)$ implies $x_{j(i)}^i \to x$.
DL2: If $x_j^i \to x_i$ for every positive integer i, and $x_{j(i)}^i \to x$ for every $j(i)$, then $x_i \to x$.

The following result is known[14]

THEOREM 12: *A necessary and sufficient condition that a D-space be a perfectly*

[13] Not to be confused with H. Kneser's "Konvergenzräume," which are convergence spaces satisfying SL1–SL3, plus a weaker condition than SL4, or with M. Fréchet's "Espaces (*L*)," which were the prototypes of all convergence spaces.

[14] Cf. "Axiomatic Definitions of Perfectly Separable Metric Spaces," by the author, Bull. Am. Math. Soc., **39** (1933), p. 607. Perfectly separable metric spaces, or "*PSM*-spaces" are precisely spaces homeomorphic with a subset of Hilbert space.

separable metric space is that it contain enumerable open sets, of which every open set is the sum of a subclass.

By an S-group (S-groupoid) is meant any continuous group (groupoid) whose elements satisfy SL1–SL4. Similarly, by a D-group (D-groupoid) is meant any S-group (S-groupoid) whose elements satisfy DL1–DL2.

11. **Transformation Convergence.** Let $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be any two convergence spaces, and consider $(1 \to 1)$ transformations σ, σ_k, etc., of $\mathfrak{S}_1$ into $\mathfrak{S}_2$. The sequence $\{\sigma_k\}$ will be said[15] to *A-converge* to σ (in symbols, $\sigma_k \to_A \sigma$) if and only if $x \subset \mathfrak{S}_1$ implies $\sigma_k(x) \to \sigma(x)$ in $\mathfrak{S}_2$.

$\{\sigma_k\}$ will be said to *B-converge* to σ (in symbols, $\sigma_k \to_B \sigma$) if and only if $x_k \to x$ in $\mathfrak{S}_1$ implies $\sigma_k(x_k) \to \sigma(x)$ in $\mathfrak{S}_2$. And $\{\sigma_k\}$ will be said to *C-converge* to σ (in symbols, $\sigma_k \to_C \sigma$) if and only if $x_k \to x$ in $\mathfrak{S}_1$ both implies and is implied by $\sigma_k(x_k) \to \sigma(x)$ in $\mathfrak{S}_2$.

The three definitions of convergence just stated ascribe to any convergence space three (in general, topologically distinct) convergence spaces. The points of all three correspond $(1,1)$ with the elements of the groupoid of the $(1 \to 1)$ transformations of the space into itself, but the definitions of L1 differ.

12. **Conditions** SL1 **and** SL2. In order that condition SL1 should hold for A-convergence, it is sufficient to restrict our attention to $(1 \to 1)$ transformations, and demand that SL1 hold in $\mathfrak{S}_2$. Since this condition is also in large measure necessary, it is suggested that the proper domain in which to use A-convergence is the domain of $(1 \to 1)$ transformations.

If we restrict ourselves to *continuous* $(1 \to 1)$ transformations, then by definition $\sigma_k = \sigma$ implies $\sigma_k \to_B \sigma$. If further we restrict ourselves to $(1,1)$ *bicontinuous* transformations, then by definition $\sigma_k = \sigma$ implies $\sigma_k \to_C \sigma$. Since these conditions are respectively necessary as well, it is suggested that B-convergence has for its proper domain of application, continuous $(1 \to 1)$ transformations, and C-convergence, bicontinuous $(1,1)$ transformations.

Moreover since A-convergence of $(1 \to 1)$ transformations of $\mathfrak{S}_1$ into $\mathfrak{S}_2$ means simply A-convergence in the transformations of the individual points of $\mathfrak{S}_1$, we see that if any of the conditions SL1–SL4 or DL2 holds in $\mathfrak{S}_2$, then it also holds in the set of the $(1 \to 1)$ transformations of $\mathfrak{S}_1$ into $\mathfrak{S}_2$. The only reason why this is not true of DL1 is that a question of uniformity enters.

Finally, if SL1 holds in $\mathfrak{S}_1$, since C-convergence and B-convergence both imply A-convergence, a fortiori[16] SL2 on $\mathfrak{S}_2$ implies that SL2 holds for the set

[15] A-convergence is the definition treated by the author in greater detail in an article "Hausdorff Groupoids," *Annals of Mathematics*, Vol. 35 (1934), pp. 351–360. B-convergence is the definition used by H. Kneser and R. Baer in articles already cited. It is also the definition of *isotopy*, as stated, for instance, in S. Lefschetz's "Topology," New York, 1930, p. 77. I am unaware of previous researches into the properties of C-convergence.

[16] For $\sigma_k \to_C \sigma$ implies $\sigma_k \to_B \sigma$ implies $\sigma_k \to_A \sigma$, and $\sigma_k \to_C \tau$ implies $\sigma_k \to_B \tau$ implies $\sigma_k \to_A \tau$. And from $\sigma_k \to_A \sigma$ and $\sigma_k \to_A \tau$ follows $\sigma = \tau$.

of the $(1 \to 1)$ transformations of $\mathfrak{S}_1$ into $\mathfrak{S}_2$, relative both to B- and to C-convergence.

13. SL3 **and** SL4. Suppose $\sigma_k \to_C \sigma$; then obviously $\sigma_k \to_B \sigma$. Further, since $\sigma_k(x) \to \sigma(x)$ implies $x_k \to x$, by defining $y_k = \sigma_k(x_k)$ and $y = \sigma(x)$, we see that if the σ_k and σ are $(1,1)$, then $y_k \to y$ implies $\sigma_k^{-1}(y_k) \to \sigma(y)$, whence $\sigma_k^{-1} \to_B \sigma^{-1}$. Conversely, if $\sigma_k \to_B \sigma$ and $\sigma_k^{-1} \to_B \sigma^{-1}$, and we consider only $(1,1)$ transformations, not only does $x_k \to x$ imply $\sigma_k(x_k) \to \sigma(x)$, but $\sigma_k(x_k) = y_k \to y = \sigma(x)$ implies $x_k = \sigma_k^{-1}(y_k) \to \sigma^{-1}(y) = x$ as well. Or

THEOREM 13: *For* $(1,1)$ *transformations,* $\sigma_k \to_C \sigma$ *is equivalent to* $\sigma_k \to_B \sigma$ *plus* $\sigma_k^{-1} \to_B \sigma^{-1}$.

COROLLARY: $\sigma_k \to_C \sigma$ *implies* $\sigma_k^{-1} \to_C \sigma^{-1}$.

Let us now return to the more general case of $(1 \to 1)$ transformations σ, σ_k, etc., of $\mathfrak{S}_1$ into $\mathfrak{S}_2$.

Suppose $\sigma_k \to_B \sigma$, and $k(i) \to \infty$. If $y_i \to y$, setting $x_{k(i)} = y_i$ and $x_k = y$ for $k \neq k(i)$, we see that if SL1 and SL4 hold in $\mathfrak{S}_1$, then $x_k \to y$, whence $\sigma_k(x_k) \to \sigma(y)$. If in addition SL3 holds in $\mathfrak{S}_2$, then $\sigma_{k(i)}(y_i) \to \sigma(y)$, whence $\sigma_{k(i)} \to_B \sigma$. It follows from Theorem 13 that if SL1, SL3, and SL4 hold in both the $\mathfrak{S}_i$, then $\sigma_k \to_C \sigma$ and $k(i) \to \infty$ imply $\sigma_{k(i)} \to_C \sigma$ for $(1,1)$ transformations.

Again, let σ_k be given, and suppose that to any $k(i) \to \infty$ corresponds $n(i) \to \infty$ satisfying $\sigma_{k(n(i))} \to_B \sigma$. Then if $x_k \to x$ and SL3 holds in $\mathfrak{S}_1$, given $k(i) \to \infty$, since $x_{k(n(i))} \to x$ for any, and $\sigma_{k(n(i))} \to \sigma$ for suitable, $n(i) \to \infty$, we can infer that for suitable $n(i) \to \infty$, $\sigma_{k(n(i))}(x_{k(n(i))}) \to \sigma(x)$. Hence if in addition SL4 holds in $\mathfrak{S}_2$, then $\sigma_k(x_k) \to \sigma(x)$, i.e., $\sigma_k \to_B \sigma$. It follows from Theorem 13 that if SL3 and SL4 hold in both the $\mathfrak{S}_i$, while to every $k(i) \to \infty$ corresponds $n(i) \to \infty$ such that $\sigma_{k(n(i))} \to_C \sigma$, then (again for $(1,1)$ transformations) $\sigma_k \to_C \sigma$.

We can summarize some of the results of §§12–13 in

THEOREM 14: *If* $\mathfrak{S}_1$ *and* $\mathfrak{S}_2$ *are S-spaces, then the* $(1 \to 1)$ *transformations of* $\mathfrak{S}_1$ *into* $\mathfrak{S}_2$ *constitute an S-space relative to A-convergence, the* $(1 \to 1)$ *continuous transformations do relative to A- or B-convergence, and the* $(1,1)$ *bicontinuous transformations do relative to A-, B-, or C-convergence.*

14. **Condition** DL2. We shall see in Theorem 25 that transformations between D-spaces do not always constitute D-spaces.[17] In spite of this, we can easily prove

THEOREM 15: *If* $\mathfrak{S}_1$ *satisfies* SL1 *and* $\mathfrak{S}_2$ *satisfies* DL2, *then the* $(1 \to 1)$ *transformations of* $\mathfrak{S}_1$ *into* $\mathfrak{S}_2$ *satisfy* DL2 *relative to B-convergence.*

For suppose $\tau_j^i \to_B \tau_i$, while $\tau_{j(i)}^i \to_B \tau$. Then, given $x_i \to x$, by SL1 $\tau_j^i(x_i) \to \tau_i(x_i)$, while certainly $\tau_{j(i)}^i(x_i) \to \tau(x)$. Consequently by DL2 $\tau_i(x_i) \to \tau(x)$, and by definition $\tau_i \to_B \tau$.

[17] That is, even the $(1,1)$ bicontinuous transformations of a Hausdorff space into itself are not a T-group in the sense of van Dantzig ("Zur topologische Algebra. I. Komplettierungstheorie," Math. Ann. **107** (1933), 587–626). It is for this reason that Hausdorff's Axioms are not suitable for the study of the topology of sets of transformations.

Corollary: *If* $\mathfrak{S}_1$ *and* $\mathfrak{S}_2$ *satisfy* SL1 *and* DL2, *then the* (1,1) *transformations of* $\mathfrak{S}_1$ *into* $\mathfrak{S}_2$ *satisfy* DL2 *relative to C-convergence.*

15. **Transformation** S**-groups.** Let $\mathfrak{S}_1$, $\mathfrak{S}_2$ and $\mathfrak{S}_3$ be any three convergence spaces, let σ_k, σ denote $(1 \to 1)$ transformations of $\mathfrak{S}_1$ into $\mathfrak{S}_2$, and τ_k, τ, $(1 \to 1)$ transformations of $\mathfrak{S}_2$ into $\mathfrak{S}_3$. Suppose $\sigma_k \to_B \sigma$ and $\tau_k \to_B \tau$.

Given $x_k \to x$ in $\mathfrak{S}_1$, we know by definition of $\sigma_k \to_B \sigma$ that $y_k = \sigma_k(x_k) \to \sigma(x) = y$, in $\mathfrak{S}_2$. Hence, since $\tau_k \to_B \tau$, $\sigma_k\tau_k(x_k) = \tau_k(y_k) \to \tau(y) = \sigma\tau(x)$. That is, $\sigma_k\tau_k \to \sigma\tau$, and, combining with Theorems 10 and 14, we get[18]

Theorem 16: *The* $(1 \to 1)$ *continuous transformations of any S-space into itself constitute an S-groupoid relative to B-convergence.*

And from Theorems 11, 13, 14, and 16 we see

Theorem 17: *The* (1,1) *bicontinuous transformations of any S-space into itself constitute an S-group under C-convergence.*

Theorem 17 is to be regarded as the backbone of the remainder of this paper; that is, our main interest will be in the nature of the S-group of the (1,1) bicontinuous transformations of given S-spaces. But our first concern will be, in §§16–20, to see under what circumstances, if any, the more complicated definition of C-convergence can be replaced by the simpler definitions of A- or B-convergence.

16. A**-convergence Implies** C**-convergence for Lines.** Let us examine the implications of A-convergence with respect to (1,1) bicontinuous transformations τ, τ_k, etc., of the line segment $\mathfrak{L}$: $0 \leqq x \leqq 1$ of the x-axis into itself—or, as we shall say for brevity, with respect to "shifts" of $\mathfrak{L}$.

Suppose $\tau_k \to_A \tau$. $\tau_k(x)$ is certainly monotonic, and evidently increasing or decreasing according as $\tau(x)$ is, for sufficiently large k. Let us suppose *both* to be increasing.

Given $\epsilon > 0$ and x_0, then for $k > k(x_0, \epsilon)$ and any x satisfying $x_0 - \epsilon < x < x_0 + \epsilon$,

$$\tau(x_0 - \epsilon) - \epsilon < \tau_k(x_0 - \epsilon) < \tau_k(x) < \tau_k(x_0 + \epsilon) < \tau(x_0 + \epsilon) + \epsilon .$$

Hence if $x_k \to x$, $\tau_k(x_k) \to \tau(x)$.

But it will be shown in Theorem 20 that in this case $\tau_k(x_k) \to \tau(x)$ implies $x_k \to x$, proving

Theorem 18: *Relative to the shifts of a line segment, the definitions of A-, B-, and C-convergence are effectively identical.*

It is pretty obvious that Theorem 18 can be generalized to include the shifts of any finite class of line segments. That it cannot be generalized much further, will be shown in §17.

17. **Inadequacy.** Let POQ be any open angle in the plane, and let PQ, PQ',

[18] This piece of reasoning essentially duplicates Kneser's (op. cit.).

and PQ'' be any three arcs connecting P with Q subject to the conditions (i) no ray through O cuts any of them more than once, and (ii) each lies within the last (fig. 1).

Now let $\overrightarrow{OX}$ be any ray from O lying in the angle POQ. Let R be the intersection of $\overrightarrow{OX}$ and PQ''; S that of $\overrightarrow{OX}$ and PQ', and T that of $\overrightarrow{OX}$ and PQ. By a "focal shift" of the interior of the figure bounded by $\overline{OP}$, PQ, and $\overline{QO}$ is meant the transformation which carries the segment $\overline{OS}$ proportionally into $\overline{OR}$, and $\overline{ST}$ proportionally into $\overline{RT}$. A focal shift evidently leaves the boundary of the figure fixed, while it transforms PQ' into PQ''.

Now for a concrete example[19] restricting the generalization of Theorem 18. Let $\mathfrak{S}$ be any manifold containing a two-dimensional region, which we may represent as the right isosceles triangle $x \leqq 1, x - y \geqq 0, x + y \geqq 0$ in Cartesian

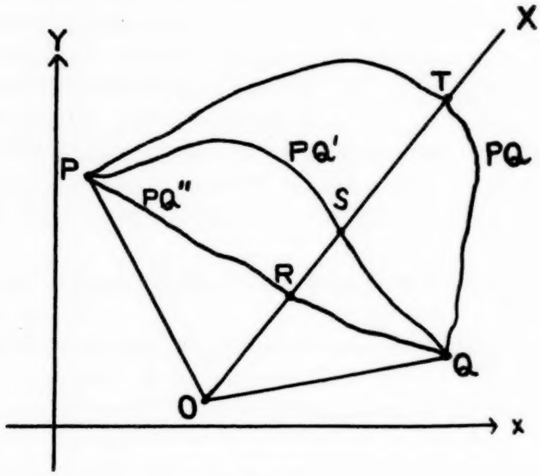

FIG. 1

2-space. Let τ_k leave all points of $\mathfrak{S}$ exterior to the subtriangle W_k bounded by the vertices $(0,0)$, $(1,1/k)$, $(1,1/(k+1))$ fixed. Let it perform a *focal shift* on W_k, in which PQ' is taken as the broken line $(1,1/k)$ to $(\frac{1}{2},1/(2k+1))$ to $(1,1/(k+1))$, and PQ'' as the broken line $(1,1/k)$ to $(\frac{1}{4},1/(4k+2))$ to $(1,1/(k+1))$. Then $\tau_k \rightarrow_A 1$.

Again, let σ_k $[k = 2, 3, 4, \cdots]$ be the focal shift with $(1,1)$ as O, $(0,0)$ as P, $(1,0)$ as Q, the circle-arc with $(\frac{1}{2},k)$ for center as PQ, the broken line $(0,0)$ to $(\frac{1}{2},1/(2k+1))$ to $(1,0)$ as PQ', and the broken line $(0,0)$ to $(\frac{1}{2},2/(2k+1))$ to $(1,0)$ as PQ''. Evidently also $\sigma_k \rightarrow_A 1$.

The reader will be able to visualize the transformations if he draws a figure analogous to figure one.

[19] This is in its underlying principles the same as the classical construction discriminating between convergence and uniform convergence.

But $\sigma_k \tau_k \nrightarrow_A 1$, because $\sigma_k \tau_k(p) \to p$ for no p: $(x,0)$ satisfying $\frac{1}{4} < x < \frac{1}{2}$. This shows

THEOREM 19: *If* $\mathfrak{S}$ *contains a two-dimensional region, then its shifts do not satisfy* LG1 *relative to A-convergence.*

Comparing with Theorems 16 and 17, we get the

COROLLARY: *If* $\mathfrak{S}$ *contains a two-dimensional region, then A-convergence implies neither B- nor C-convergence.*

Since the construction proving Theorem 19 can be duplicated in its essential features in case $\mathfrak{S}$ contains an n-dimensional region,[20] we see that Theorem 18 cannot have any very important generalizations.

18. *B*-convergence Often Implies *C*-convergence. *C*-convergence always implies *B*-convergence; we shall see that the two definitions are effectively equivalent for two important types of spaces. For instance, we can prove

THEOREM 20: *B-convergence and C-convergence are effectively equivalent for* (1,1) *bicontinuous transformations between compact S-spaces.*

Suppose (denoting by τ, τ_k, etc., (1,1) bicontinuous transformations of the compact S-space $\mathfrak{S}_1$ into the compact S-space $\mathfrak{S}_2$) $\tau_k \to_B \tau$, while $\tau_k(x_k) \to \tau(x)$. Since $\mathfrak{S}_2$ is compact, to any $k(i) \to \infty$ corresponds $n(i) \to \infty$ such that $x_{k(n(i))} \to y$; hence by Theorem 13 and SL3 $\tau_{k(n(i))}(x_{k(n(i))}) \to \tau(y)$, and by SL2 $\tau(y) = \tau(x)$. Since τ is (1,1), we can infer that $y = x$. Hence, by SL4, $x_k \to x$, and by definition $\tau_k \to_C \tau$.

COROLLARY: *The shifts of any compact S-space constitute an S-group relative to B-convergence.*

We shall digress to point out that there exists a compact "Konvergenzraum"[21] $\mathfrak{S}$, whose shifts do not satisfy LG2 relative to B-convergence (Kneser's convergence definition).

Let the points of $\mathfrak{S}$ be x_k $[k = 1, 2, 3, \cdots]$ and x. Define $x_{k(i)} \to x$ *if* $k(i) \to \infty$, *unless* there are an infinite number of odd integers $(2n + 1)$ such that $k(i) = 2n + 1$ has more than one solution.

Define the sequence $\tau_1, \tau_2, \tau_3, \cdots$ of shifts of $\mathfrak{S}$ as follows,

$$\tau_{2k}\colon x_{2k+1} \to x_{4k},\ x_{2k+1+2i} \to x_{2k-1+2i},\ x_{4k+2i-2} \to x_{4k+2i},$$
$$[i = 1, 2, 3, \cdots];\ \text{the other } x_i \text{ stay fixed.}$$
$$\tau_{2k+1}\colon x_{2k+1} \to x_{4k+2},\ x_{2k+1+2i} \to x_{2k-1+2i},\ x_{4k+2i} \to x_{4k+2i+2},$$
$$[i = 1, 2, 3, \cdots];\ \text{the other } x_i \text{ stay fixed.}$$

Then $\tau_k \to_B 1$, yet since $\tau_k^{-1}(x_{2k}) \nrightarrow x$, $\tau_k^{-1} \nrightarrow_B 1$.

19. Second Example. Suppose $\mathfrak{S} = C_1 + C_2 + C_3 + \cdots$, where (i) $C_k \subset C_{k+1}$, (ii) each C_k is a compact connected set, (iii) C_k contains no point or

[20] By "region" we mean of course open subset, and by n-dimensional region, one homeomorphic with a region in Cartesian n-space; such a one will contain a subregion homeomorphic with an n-simplex.

[21] H. Kneser, op. cit., p. 365.

limit point of $\mathfrak{S} - C_{k+1}$, and (iv) $\mathfrak{S}$ is a D-space. Any such space $\mathfrak{S}$ will be called an *O-space*. For instance, any open connected set in Cartesian n-space is an O-space.

THEOREM 21: *For shifts of any O-space $\mathfrak{S}$, B-convergence and C-convergence are effectively equivalent.*

Suppose $\tau_k \to_B \tau$, and $\tau_k(x_k) \to \tau(x)$. By SL1 and $\tau_k \to_B \tau$, $\tau_k(x) \to \tau(x)$. Let C_T be the first C_k to contain $\tau(x)$. By (iii), if we omit a finite number of the k, $\tau_k(x_k) \subset C_{T+1}$ and $\tau_k(x) \cdot \subset C_{T+1}$. Hence by (ii) we can join $\tau_k(x)$ and $\tau_k(x_k)$ by a curve Γ_k lying wholly in C_{T+1}.

$\tau^{-1}(C_{T+1})$, being compact, by (iii) lies wholly in some C_J. And unless $\tau_k^{-1}(\Gamma_k) = \Delta_k$ lies wholly in C_{J+1}, we can choose $y_k \subset \Delta_k$ on the boundary of C_{J+1}. Hence if enumerable $\Delta_k \not\subset C_{J+1}$, by (ii) there exists a subsequence of such points $y_{k(i)}$, converging to some limit point y, which will by DL1 be a boundary point of C_{J+1}. But because $\tau_{k(i)}(y_{k(i)}) \subset \Gamma_{k(i)} \subset C_{T+1}$, we know that $\tau_{k(i)}(y_{k(i)}) \to \tau(y) \subset C_{T+1}$, and $y \subset \tau^{-1}(C_{T+1}) \subset C_J$. This contradicts (iii), since y is by hypothesis a limit point of $\mathfrak{S} - C_{J+1}$.

Therefore all but finite $\Delta_k \subset C_{J+1}$. And from here we can proceed (C_{J+1} being compact) as in the proof of Theorem 20.

It is evident that the conclusion of Theorem 21 can be extended to the sum of a finite number of mutually separated O-spaces; it is at least a plausible conjecture that it is true for all homogeneous sets in Cartesian space.

20. **Inadequacy.** Unfortunately, there is an important class of spaces whose shifts do not constitute a continuous group relative to B-convergence. We shall cite a simple example, and leave it to the reader to generalize the construction.

THEOREM 22: *There is a plane set $\mathfrak{S}$ whose shifts do not satisfy* LG2 *relative to B-convergence.*

$\mathfrak{S}$ contains $(0,0)$, and in addition the interior of the triangle whose vertices are $(-1,0)$, $(1,1)$, and $(1,0)$—using Cartesian coordinates. τ_k leaves the exterior of the triangle W_k whose vertices are $(-1,0)$, $(1/k, 1/k + 4/3k^2)$, and $(-\frac{1}{2},0)$ fixed, performs a focal shift (cf. §17) on W_k, relative to the broken lines $(-1,0)$ to $(0,1/k)$ to $(-\frac{1}{2},0)$, and $(-1,0)$ to $(-\frac{1}{4},2/3k)$ to $(-\frac{1}{2},0)$, and to the "focus" $(1/k, 1/k + 4/3k^2)$.

The reader will be able to visualize the transformations if he draws a figure analogous to figure one.

Suppose $p_k \to p$. If $p \neq (0,0)$, then the y-coordinate of p is positive, and for k large enough, since $p_k \not\subset W_k$, we have $\tau_k(p_k) = p_k \to p = 1(p)$. While if $p = (0,0)$, then since τ_k *increases* the distance from p of no point by more than $1/k$, it is still true that $\tau_k(p_k) \to p$. Hence by definition $\tau_k \to_B 1$.

Now set $q_k = (0,1/k)$, so that $\tau_k^{-1}(q_k) = (-\frac{1}{4},2/3k)$. Although $q_k \to q = (0,0)$, $\tau_k^{-1}(q_k) \not\to q$; hence $\tau_k^{-1} \not\to_B 1$, and the theorem is proved.[22]

[22] This example contradicts the principle asserted by H. Kneser in footnote 3, p. 366, Math. Zeits., 25 (1926). But our Theorems 20 and 21 confirm the applications he makes of the principle.

21. **The Shifts of Compact Metric Spaces.** Theorems 17 and 22 show that in general we must regard C-convergence as fundamental, if we wish to incorporate the ideas of Schreier into the theory of the topology of transformation-sets. We shall now show that for important classes of metric spaces, we can include in the same definitions the theory of the topology of Hilbert space.

A metric space is of course one in which to every pair (x, y) of points corresponds a number $\overline{xy}$ called the "distance" between x and y, satisfying

M1: $\overline{xx} = 0$; if $x \neq y$, then $\overline{xy} > 0$.
M2: $\overline{xy} = \overline{yx}$ irrespective of x and y.
M3: $\overline{xy} \leqq \overline{xz} + \overline{zy}$ irrespective of x, y, and z.

Theorem 23: *If $\mathfrak{S}$ is a compact metric space, then the (continuous) group G of its shifts relative to C-convergence is homeomorphic with a perfectly separable metric space.*

Let σ and τ be any two shifts of $\mathfrak{S}$. By $\overline{\sigma\tau}$ we mean $\overline{\mathrm{Lim}}_{x \subset \mathfrak{S}}\, \overline{\sigma(x)\tau(x)}$. We shall prove the theorem in three parts: (i) $\sigma_k \to_C \sigma$ if and only if $\overline{\sigma\sigma_k} \to 0$ (ii) the distance function $\overline{\sigma\tau}$ is metric (iii) the space of the shifts is perfectly separable.

$\overline{\sigma(x_k)\sigma_k(x_k)} \leqq \overline{\sigma\sigma_k}$. And since σ is continuous, $\overline{xx_k} \to 0$ implies $\overline{\sigma(x)\sigma(x_k)} \to 0$. Hence if $\overline{\sigma\sigma_k} \to 0$ $\overline{xx_k} \to 0$, then $\overline{\sigma(x)\sigma_k(x_k)} \leqq \overline{\sigma(x)\sigma(x_k)} + \overline{\sigma(x_k)\sigma_k(x_k)} \to 0$, and $\sigma_k \to_B \sigma$. Conversely if $\overline{\sigma\sigma_k} \nrightarrow 0$, we can choose a subsequence $\tau_i = \sigma_{k(i)}$ such that $\overline{\sigma\tau_i} \geqq \epsilon > 0$, and x_i such that $\overline{\sigma(x_i)\tau_i(x_i)} > \frac{3}{4}\,\epsilon$. Since $\mathfrak{S}$ is compact, there exists a subsequence $y_j = x_{i(j)}$ converging to some point y; setting $\rho_j = \sigma_{i(j)}$ we get $\overline{\sigma(y)\sigma(y_j)} \to 0$ and consequently $\overline{\sigma(y)\rho_j(y_j)} \geqq \overline{\sigma(y_j)\rho_j(y_j)} - \overline{\sigma(y)\sigma(y_j)} > \frac{1}{2}\,\epsilon$ for sufficiently large j. That is, $\sigma_k \nrightarrow_B \sigma$, whence we see that $\overline{\sigma\sigma_k} \to 0$ and $\sigma_k \to_B \sigma$ are equivalent. Using Theorem 20, we get (i).

Obviously $\overline{\sigma\tau} = 0$ if $\sigma = \tau$; $\overline{\sigma\tau} > 0$ if $\sigma \neq \tau$. Equally obviously $\overline{\sigma\tau} = \overline{\tau\sigma}$. We can also easily prove M3, since for arbitrarily small positive ϵ, $\overline{\sigma(x)\tau(x)} = \overline{\sigma\tau} - \epsilon$, whence $\overline{\sigma\rho} + \overline{\rho\tau} \geqq \overline{\sigma(x)\rho(x)} + \overline{\rho(x)\tau(x)} \geqq \overline{\sigma(x)\tau(x)} \geqq \overline{\sigma\tau} - \epsilon$. This proves (ii).

Let $\epsilon > 0$ and τ be given. Set $\delta(\epsilon, \tau) = \underline{\mathrm{Lim}}_{\overline{\tau(x)\tau(y)} \geqq \epsilon}\, \overline{xy}$. Since $\mathfrak{S}$ is compact, $\delta(\epsilon, \tau) > 0$ identically; otherwise we could find a limit point x and sequences $\{x_k\}$ and $\{y_k\}$ such that $\overline{xx_k} \to 0$ and $\overline{x_ky_k} \to 0$, yet $\overline{\tau(x_k)\tau(y_k)} \geqq \epsilon$. There would ensue $\overline{xy_k} \leqq \overline{xx_k} + \overline{x_ky_k} \to 0$, therefore $\overline{\tau(x)\tau(x_k)} \to 0$ and $\overline{\tau(x)\tau(y_k)} \to 0$, contradicting $\epsilon \leqq \overline{\tau(x_k)\tau(y_k)} \leqq \overline{\tau(x_k)\tau(x)} + \overline{\tau(x)\tau(y_k)}$. Accordingly, if R_k is the domain of the τ for which $\delta(\epsilon, \tau) \geqq 1/k$, then

$$R_1 \subset R_2 \subset R_3 \subset \cdots \quad \text{and} \quad R_1 + R_2 + R_3 + \cdots = G\,.$$

Let us fix on a particular R_k, and choose in it successively $\tau_1, \tau_2, \tau_3, \cdots$ as long as possible, subject to the condition $i < n + 1$ implies $\overline{\tau_i\tau_{n+1}} \geqq 5\epsilon$. Suppose we got in this way an infinite sequence of τ_i.

For each j, x_j' would exist satisfying $\overline{\tau_1(x_j')\tau_j(x_j')} \geqq 4\epsilon$. Moreover $\{x_j'\}$ would have a limit point x_1, to which would correspond an infinite subsequence of $\{x_j'\}$ satisfying $\overline{x_1x_j'} < 1/k$. Consequently

$$\begin{aligned}\overline{\tau_1(x_1)\tau_j(x_1)} &\geqq \overline{\tau_1(x_j')\tau_j(x_j')} - \overline{\tau_1(x_1)\tau_1(x_j')} - \overline{\tau_j(x_j')\tau_j(x_1)} \\ &\geqq 4\epsilon - \epsilon - \epsilon = 2\epsilon\,.\end{aligned}$$

Proceeding inductively, we could get $\tau_1, \tau_2, \tau_3, \cdots$ and $x_1, x_2, x_3, \cdots$ such that $i < j$ implied $\overline{\tau_i(x_i)\tau_j(x_i)} \geqq 2\epsilon$. Since $\mathfrak{S}$ is compact, $\{x_i\}$ would have a limit point y, and a subsequence $\{y_k\}$ which converged to y. Restricting to $\overline{yy_k} < 1/2k$, whence $\overline{y_iy_j} < 1/k$, we would have a new infinite subsequence $\{z_k\}$ such that $i < j$ implied $\overline{\tau_i(z_i)\tau_j(z_j)} \geqq \overline{\tau_i(z_i)\tau_j(z_i)} - \overline{\tau_j(z_i)\tau_j(z_j)} = 2\epsilon - \epsilon = \epsilon$. Therefore the $\tau_k(z_k)$ could not have a limit point, contradicting the compactness of $\mathfrak{S}$.

That is, we obtain a finite sequence $\tau_1, \cdots, \tau_n$ such that every τ of R_k satisfies $\overline{\tau\tau_i} < 5\epsilon$ for some i; that is, the τ_i are 5ϵ-dense in R_k. Summing for k, and again for $\epsilon = \frac{1}{2}, \frac{1}{4}, 1/8, \cdots$ we have an enumerable set of τ_i dense in G, giving us[23] (iii).

22. **Generalization.** We can make the following important generalization of Theorem 23, namely

THEOREM 24: *If* $\mathfrak{S} = C_1 + C_2 + C_3 + \cdots$, *and* (i) $C_k \subset C_{k+1}$ (ii) *the* C_k *are compact metric sets* (iii) $x_k \not\subset C_k$ *precludes the convergence of* $\{x_k\}$, *then the group* G *of the shifts of* $\mathfrak{G}$ *is homeomorphic (under C-convergence) with a subset of Hilbert space.*

Set $r_n(\sigma, \tau) = \text{Min}\,\{1, \overline{\text{Lim}}_{x \subset c_n}\, (\overline{\sigma(x)\tau(x)} + \overline{\sigma^{-1}(x)\tau^{-1}(x)})\}$. Then define $\overline{\sigma\tau} = \sum_{n=1}^{\infty} (1/n^2) \cdot r_n(\sigma, \tau)$. Evidently $\overline{\sigma\tau} = \overline{\sigma^{-1}\tau^{-1}}$, and we can show as in Theorem 23 that $\overline{\sigma\tau}$ satisfies M1–M3.

Suppose $\overline{\sigma\sigma_k} \to 0$, and let $x_k \to x$ be given. By (iii) and SL3 we can choose N so large that $C_N \supset \{x_k\} + x$. So soon as $\overline{\sigma\sigma_k} < \epsilon/N^2$, and $\overline{\sigma(x)\sigma(x_k)} < \epsilon$, then $\overline{\sigma(x)\sigma_k(x_k)} < 2\epsilon$; similarly $\overline{\sigma^{-1}(x)\sigma_k^{-1}(x_k)} < 2\epsilon$. Therefore $\overline{\sigma\sigma_k} \to 0$ implies $\sigma_k \to_C \sigma$.

Conversely, suppose $\overline{\sigma\sigma_k} \nrightarrow 0$. We can choose a subsequence $\sigma_{k(i)}$ satisfying $\overline{\sigma\sigma_{k(i)}} > 3/L$, and hence, since $\sum_{n=L+1}^{\infty} r_n(\sigma, \tau) \cdot (1/n^2) < 1/L$, $x_i \subset C_L$ satisfying $\overline{\sigma(x_i)\sigma_i(x_i)} > 1/L$ or else $\overline{\sigma^{-1}(x_i)\sigma^{-1}(x_i)} > 1/L$. In any case, by taking a convergent subsequence, we can prove (essentially as in Theorem 23) that $\sigma_k \nrightarrow_C \sigma$. This proves that G is homeomorphic (under C-convergence) with the metric space defined in paragraph one. It remains to show that an everywhere dense enumerable set exists in G.

To show this, observe that, given (i, j), we can prove (again as in Theorem

[23] Cf. for example S. Lefschetz, "Topology," New York, 1930, p. 6.

23) that there are enumerable transformations dense with respect to transformations τ such that $\tau(C_i) \subset C_j$ while $\tau^{-1}(C_i) \subset (C_j)$. But the (enumerable) sum of these is dense in G, as the reader can easily check.

COROLLARY: *In Theorems* 23 *and* 24, *G is a T-group in the sense of van Dantzig.*[24]

In view of the intimate connection between continuous transformations and continuous functions on the one hand, and continuous functions and Hilbert space on the other, we cannot regard Theorems 23 and 24 as very surprising; what is more surprising is that they are not susceptible of more universal generalizations.

23. **Restriction on generality.** In fact, it is not even true that shift-groups of simple metric spaces are in general T-groups in the sense of van Dantzig. We shall illustrate this by an example, which the reader will see is capable of wide generalization.

Let $\mathfrak{S}$ be the sum of the point $(0,0)$ and the rectangle interior $-1 < x < 1$, $0 < y < 2$. Let τ_j^i $[i = 3, 4, 5, \cdots; j = 1, 2, 3, \cdots]$ leave the exterior of the triangle W_j whose vertices are p_j: $(-1, 1/j + 1/j^2)$, q_j: $(0, 1/j)$, and r_j: $(-1, 1/j - 1/j^2)$ fixed; let it perform a focal shift (§17) on W_j, away from the "focus" q_j, carrying the broken line p_j to $(-1/i, 1/j)$ to r_j into the broken line p_j to $(-\frac{1}{2}, 1/j)$ to r_j.

$\tau_j^i \rightarrow_C 1 = \tau_i$ (as in the proof of Theorem 19), and consequently $\tau_i \rightarrow_C 1$. Nevertheless no matter what $N(i)$ we choose, if we set $j(i)$ greater than i and $N(i)$, then $\tau_{j(i)}^i(-1/i, 1/j(i)) = (-\frac{1}{2}, 1/j(i)) \nrightarrow (0,0)$, in spite of the fact that $(-1/i, 1/j(i)) \rightarrow (0,0)$. This proves

THEOREM 25: *There exists a plane set whose shifts do not satisfy* DL1 *relative to C-convergence.*

Theorem 25 shows why the shifts of a given space are not a Hausdorff space; any Hausdorff space satisfies DL1.

It might be thought, because of the parallelism between Theorems 17–19 and Theorems 23–25, that spaces whose shifts were homeomorphic with a subset of Hilbert space were spaces in which B-convergence was effectively identical with C-convergence. To destroy this conjecture, the plane set of points $(i, 1/j)$ and $(i, 0)$ $[i, j = 1, 2, 3, \cdots]$ is sufficient. Its shifts are homeomorphic (under C-convergence) with a subset of Hilbert space, yet in it B-convergence and C-convergence are not effectively equivalent.

24. **Linear transformations.** Linear transformations deserve a special study because of their central rôle in analysis. Of course in the case of finite dimensions, the set of the linear transformations of one vector-space into any other is homeomorphic with an easily identified subset of ordinary Cartesian space.

Concerning the more complex case of infinite dimensions, we shall prove a few basic and easily demonstrable theorems. We shall deal exclusively with

[24] "Topologische Algebra. I. Komplettierungstheorie," Math. Ann. **107** (1933), 587–626.

"B-spaces" in the sense of Banach[25]—that is, with metrized, topologically complete vector spaces in which distance is invariant under simple translations.

It is known[26] that any $(1 \to 1)$ linear (continuous) transformation τ of a first B-space $\mathfrak{B}_1$ into a second B-space $\mathfrak{B}_2$ has a "norm" $|\tau|$ with the property that $|\tau(x) - \tau(x')| \leqq |\tau| \cdot |x - x'|$ for any x and x' in $\mathfrak{B}_1$. Further, it is known[27] that if $\tau_n \to_A \tau$, then $|\tau_n|$ is bounded, say by τ^*.

In virtue of the simple formula

$$\begin{aligned} |\tau_n(x_n) - \tau(x)| &\leqq |\tau_n(x) - \tau_n(x)| + |\tau_n(x) - \tau(x)| \\ &\leqq \tau^* \cdot |x_n - x| + |\tau_n(x) - \tau(x)| \end{aligned}$$

we can say that if also $|x_n - x| \to 0$, then $|\tau_n(x_n) - \tau(x)| \to 0$. That is, $\tau_n \to_A \tau$ implies $\tau_n \to_B \tau$, and (the converse being obvious)

THEOREM 26: *A-convergence and B-convergence are effectively equivalent for linear transformations of B-spaces.*

COROLLARY: *If a sequence* $\{\sigma_n\}$ *of linear transformations of the B-space* $\mathfrak{B}_1$ *into the B-space* $\mathfrak{B}_2$ *A-converges to the limit* σ, *while the sequence* $\{\tau_n\}$ *of linear transformations of* $\mathfrak{B}_2$ *into* $\mathfrak{B}_3$ *A-converges to the limit* τ, *then* $\sigma_n \tau_n \to_A \sigma\tau$.

However, A-convergence does not imply C-convergence, even for $(1,1)$ bicontinuous transformations of the Hilbert space $\mathfrak{H}$ into itself. To see this, let $\xi_1, \xi_2, \xi_3, \cdots$ be a complete set of "orthonormal"[28] points in $\mathfrak{H}$, and let $\tau_n(\xi_k) = \xi_k$ when $k \neq n$, while $\tau_n(\xi_n) = (1/n)\cdot\xi_n$. $\tau_n \to_A 1$, yet since $|\tau_n^{-1}| = n$ is not bounded, $|\tau_n^{-1}| \not\to_A 1$.

THEOREM 27: *Neither the functionals nor the* $(1,1)$ *bicontinuous linear transformations of Hilbert space satisfy* DL1.

To prove the first half of Theorem 27, let $f_j^i(\xi_k) = 0$ if $k \neq 1$ or j; let $f_j^i(\xi_1) = 1$, and $f_j^i(\xi_j) = i$. Then for fixed i, $f_j^i \to_A f$ [where $f(\xi_1) \equiv 1$ and $f(\xi_k) \equiv 0$ if $k \neq 1$]. Yet since $|f_j^i| = i$, $f_{j(i)}^i \not\to_A f$ no matter how we choose $j(i)$.

The second half can be similarly proved by setting $\tau_j^i(\xi_k) = \xi_k$ when $k \neq j$, and $\tau_j^i(\xi_j) = i\cdot\xi_j$.

The effect of a change in the definition of the convergence of transformation sequences on the topology of transformation-sets is strikingly illustrated by the contrast between Theorem 27 and the fact that the linear transformations of a B-space into itself are again a B-space under the obvious definitions of the sum of two transformations and the product of a transformation by a constant, and Banach's definition of the "norm" of a transformation. That is, being metric, they certainly satisfy DL1 if we replace ordinary convergence by uniform convergence.

HARVARD UNIVERSITY.

[25] S. Banach, "Théorie des operations linéaires," Warsaw, 1932; cf. p. 53, p. 26. Banach defines convergence as our A-convergence.

[26] Ibid., p. 54, Theorem 1.

[27] Ibid., p. 80, Theorem 5.

[28] Cf. M. H. Stone, "Linear transformations of Hilbert space," New York, 1932, p. 7 (Definition 1.6).

Annals of Mathematics
Vol. 35, No. 4, October, 1934

GENERALIZED CLOSED MANIFOLDS IN n-SPACE[1]

By R. L. Wilder

(Received May 8, 1934)

In recent papers[2] I have characterized 2-dimensional closed manifolds in E_3[3] by considering a manifold as a point set which is a common boundary of two domains that satisfy certain conditions. The central problem of the present work is to solve the analogous problem for n-space, with, however, the emphasis put on obtaining the characterization *from the properties of one domain alone.*

That it is possible to characterize, by internal properties alone, those domains *in the plane* whose boundaries are 1-manifolds (= simple closed curves), was recognized by R. L. Moore, who proved[4] that in order that a simply connected, bounded domain D in E_2 should have a simple closed curve as boundary, it is necessary and sufficient that D should be uniformly locally connected. One of the principal results which we obtain is an analogue of this theorem for n-space (Principal Theorem C).

For the case of ordinary 3-space, where the bounding sets are 2-dimensional, no difficulty is encountered in obtaining the manifolds in the classical sense. For higher dimensions, one can hardly hope to obtain the classical manifolds, since our knowledge of them is so meager. I therefore follow the present trend on the part of topologists of working with *generalized manifolds.*[5] I was faced, early in the progress of these investigations, with the problem of discovering conditions on a point set, matching known invariants of classical manifolds, which would be equivalent to conditions applied solely to the domain bounded by the point set. The result is the definition of generalized closed n-manifold given below. Not only does this manifold behave as it should internally (e.g., have finite Betti numbers, satisfy the Poincaré duality relation), but we

---

[1] The portion of this paper relating to 3-space and the classical 2-dimensional manifolds was presented to the Amer. Math. Soc. Dec. 27, 1933, (Bull. Amer. Math. Soc., **39** (1933), p. 875). The remainder was presented at the meeting of March 30, 1934 (ibid., **40**, p. 214).

[2] *A converse of the Jordan-Brouwer separation theorem in three dimensions*, Trans. Amer. Math. Soc., **32** (1930), pp. 632–657; *On the properties of domains and their boundaries in E_n*, Math. Ann., **109** (1933), pp. 273–306; also Theorem 8 of *Concerning a problem of K. Borsuk*, Fund. Math., **21** (1933), pp. 156–167. These papers will be denoted hereafter by I, II and III.

[3] We use the symbol E_n to denote n-dimensional euclidean space, instead of the R_n often employed.

[4] *A characterization of Jordan regions by properties having no reference to their boundaries*, Proc. Nat. Acad. Sci., **4** (1918), pp. 364–370.

[5] See S. Lefschetz, *On generalized manifolds*, Amer. Jour. Math., **55** (1933), pp. 469–504; E. Čech, *Théorie générale des variétés et de leurs théorèmes de dualité*, Annals of Math., **34** (1933), pp. 621–730.

find that its relations to the complementary space, when it occurs in E_{n+1}, are those that a classical manifold satisfies.

In §1, I show that such an $(n - 1)$-manifold in E_n is the common boundary of just two domains each of which is uniformly locally connected for dimensions from zero to $n - 2$, inclusive, in the sense defined below; and is, moreover, characterized by this condition. In §3, we see, however, that the property concerning local connectedness to which we have just referred, when applied to a single simply $(n - 1)$-connected domain, is sufficient to make the boundary of that domain a generalized closed $(n - 1)$-manifold. §2 is devoted to a study of Betti numbers and duality relations. One of the by-products of this investigation is a new duality (see Theorem 5) between the cycles of the domains bounded by an $(n - 1)$-manifold.

In §4, I derive a very general theorem (Principal Theorem E) for the characterization of manifold boundaries by conditions distributed over the two complementary domains; my earlier result (Theorem 20 of Paper II) concerning 2-manifolds in E_3 is a special case of this (with $j = 0$), since a generalized closed 2-manifold in E_3 is a manifold in the classical sense.

In conclusion, I consider in §5 the boundary of the most general bounded, open subset of E_n having the local connectedness property mentioned above. Such a point set yields natural extensions of the earlier results on manifolds. In general, it consists of a set (possibly infinite) of generalized manifolds and single points, has finite Betti numbers (for dimensions 1 to $n - 2$ inclusive) and satisfies internally the Poincaré duality relation.

§1

Preliminary definitions. Let D denote an open subset of E_n, and F denote any compact metric space. The cycles, chains, etc., of D may be based on the complexes obtained in D from a sequence of subdivisions of E_n whose meshes converge to zero.[6] For F, the base for the combinatorial procedure is the ϵ-complex in terms of which are obtained the Vietoris cycles which we employ throughout.

We call D (resp. F) *simply i-connected* (i a non-negative integer) if $p^i(D) = 0$ ($p^i(F) = 0$).

If D has the property that, given an $\epsilon > 0$, there exists a $\delta_\epsilon > 0$ such that

[6] I have tried throughout this paper to presuppose as little of set-theoretic and combinatorial notions as possible. I presuppose, on the part of the reader, some knowledge of the theory of Jordan continua, and have purposely phrased the combinatorial machinery in symbols of the modulo 2 topology. In the latter case, however, we might just as well have used any prime modulus, or rational chains—the results obtained hold just as well. It is essential, however, that a 0-cycle, in the combinatorial theory of complexes, be an even number of points. For the combinatorial framework of D, see the paper of Alexander cited in footnote 12. For the theory of the connectivity groups of a compact metric space I refer the reader to L. Vietoris, *Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen*, Math. Ann., **97** (1927), pp. 454–472.

any i-cycle of D of diameter $< \delta_\epsilon$ bounds a chain of D of diameter $< \epsilon$,[7] then we say that D is *uniformly locally i-connected* (= u.l.i-c.).

If F has the property that, given a point P and an $\epsilon > 0$ there exists a $\delta_\epsilon > 0$ such that any i-cycle of $S(P, \delta_\epsilon)$ bounds a chain in $S(P, \epsilon)$,[8] then we say that F is *locally i-connected.* (From the compactness of F it follows easily that this is a *uniform* property of F, but this does not enter into the considerations below.)

Definition M^n. Let M be a compact metric space of dimension n[9] such that 1) $p^n(M) = 1$, but if F is a closed proper subset of M, then $p^n(F) = 0$; 2) if P is a point of M, there is an $\epsilon > 0$ such that if γ^i $(1 \leqq i \leqq n - 1)$ is an i-cycle of $S(P, \epsilon)$, then $\gamma^i \sim 0$ in M;[10] 3) if P is a point of M and ϵ a positive number, there exist positive numbers δ and η, $\epsilon > \delta > \eta$, such that if $\gamma^i (0 \leqq i \leqq n - 2)$ is a cycle of $F(P, \delta)$, then $\gamma^i \sim 0$ in $S(P, \epsilon) - S(P, \eta)$; and if γ^{n-1} is a cycle of $F(P, \delta)$, then $\gamma^{n-1} \sim 0$ in $M - S(P, \eta)$. We call such a space M a *generalized closed n-dimensional manifold*; we denote it briefly by g.c.n-m.

That a domain, complementary to an $(n - 1)$-manifold in the classical sense in E_n, is uniformly locally i-connected $(0 \leqq i \leqq n - 2)$, I have shown in a previous paper.[11] This is, however, an external property of a much more general class of configurations, namely those that satisfy conditions 2) and 3) of the definition of a g.c. $(n - 1)$-m.

THEOREM 1. *In E_n, let M be a bounded, closed point set satisfying conditions 2) and 3) of definition M^{n-1}. Then if D is a domain complementary to M, D is uniformly locally i-connected* $(0 \leqq i \leqq n - 2)$.

PROOF. Consider first the case $i = 0$. Let ϵ be an arbitrary positive number and P a point of M on the boundary of D; let δ and η be numbers as provided by condition 3), and let γ^0 be a 0-cycle of $D \cdot S(P, \eta)$. There exist chains K_1^1, K_2^1 such that

$$K_1^1 \rightarrow \gamma^0 \text{ in } S(P, \eta); \text{ in } E_n - [F(P, \epsilon) + M - M \cdot S(P, \delta)]$$
$$K_2^1 \rightarrow \gamma^0 \text{ in } D; \text{ in } E_n - M \cdot \overline{S(P, \delta)} .$$

[7] By the *diameter* of a chain we mean the diameter of the minimal point set which carries the chain; for the case of an F, if $\gamma^i = \{i_1, i_2, \cdots, i_k, \cdots\}$ is a Vietoris cycle, a chain bounded by F will be simply a sequence of ϵ_k-complexes K_k^{i+1} such that $K_k^{i+1} \rightarrow i_k$ and $\lim \epsilon_k = 0$ as $k \rightarrow \infty$. See P. Alexandroff, *Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension,* Annals of Math., **30** (1928-29), pp. 101-187, especially pp. 168-169.

[8] In any metric space (usually E_n in the present paper), by $S(P, \epsilon)$ we mean the set of all points x such that the distance from x to P is $< \epsilon$; by $F(P, \epsilon)$ we mean the set of points x whose distance from P is exactly ϵ.

[9] I.e., Brouwer-Urysohn-Menger dimension.

[10] Obviously we might have said, since M is compact, that "there exists an $\epsilon > 0$ such that any γ^i of diameter $< \epsilon$ bounds on M." We prefer, however, to state the definition in the above form.

[11] Paper I, referred to in footnote [2].

We ask, does the cycle $K_1^1 + K_2^1$ link the point set

$$[F(P, \epsilon) + M - M \cdot S(P, \delta)] \cdot [M \cdot \overline{S(P, \delta)}]\,?$$

The latter set is a subset of $M \cdot F(P, \delta)$, and if $K_1^1 + K_2^1$ links $M \cdot F(P, \delta)$, it is linked with a cycle γ^{n-2} of $M \cdot F(P, \delta)$. But by condition 3), $\gamma^{n-2} \sim 0$ on $M - M \cdot S(P, \eta)$, whereas $K_1^1 + K_2^1$ meets M only in $S(P, \eta)$. Thus $K_1^1 + K_2^1$ does not link $M \cdot F(P, \delta)$, and therefore does not link the above product. Hence by the Alexander Addition Theorem,[12] $\gamma^0 \sim 0$ in $D \cdot S(P, \epsilon)$. It follows at once that D is uniformly locally 0-connected.

The case $0 < i < n - 1$ is handled in a somewhat similar fashion. First, however, if the given ϵ is not small enough we choose a smaller ϵ such that cycles of M in $\overline{S(P, \epsilon)}$ bound on M (Condition 2)). We then get δ and η as provided by condition 3), and let γ^i be a cycle of $D \cdot S(P, \eta)$. We have

$$K_1^{i+1} \to \gamma^i \text{ in } S(P, \eta);\ \text{in } E_n - [F(P, \epsilon) + M - M \cdot S(P, \delta)]\,.$$

And we observe that γ^i cannot link the point set $M \cdot \overline{S(P, \epsilon)}$, since all cycles of this set bound in M. Therefore we have

$$K_2^{i+1} \to \gamma^i \text{ in } E_n - M \cdot \overline{S(P, \epsilon)};\ \text{in } E_n - M \cdot \overline{S(P, \delta)}\,.$$

By an argument similar to the above, we see that $K_1^{i+1} + K_2^{i+2}$ does not link $M \cdot F(P, \delta)$, since all cycles γ^{n-i-2} of $M \cdot F(P, \delta)$ $(0 < i < n - 1)$ bound on $M \cdot [S(P, \epsilon) - S(P, \eta)]$ whereas K_1^{i+1} meets M only in $S(P, \eta)$ and K_2^{i+1} meets M only in $E_n - S(P, \epsilon)$.

If M is a closed, bounded subset of E_n satisfying condition 1) of definition M^{n-1}, then its complement is just two domains of which it is the common boundary, so that we have:

COROLLARY 1. *If M is a g.c. $(n - 1)$-m. in E_n, then $E_n - M$ is the sum of two uniformly locally i-connected* $(0 \leqq i \leqq n - 2)$ *domains of which M is the common boundary.*

We shall see, later on (Principal Theorem D), that if M and D are as in Theorem 1, then the boundary of D is itself a g.c. $(n - 1)$-m. which is a (in general, *proper*) subset of M.

In order to separate out the difficulties, we insert at this point a converse of Corollary 1. For the proof of this we need a lemma on intersections of chains based on complexes formed out of subdivisions of E_n.

LEMMA 1. *Let γ^i and Γ^{n-i-1} be cycles such that γ^i links Γ^{n-i-1}, and let K_1^{i+1} and K_2^{i+1} be chains bounded by γ^i. Also, let S^{i+2} be any chain bounded by $K_1^{i+1} + K_2^{i+1}$. Then some component*[13] *of the intersection, L^1, of Γ^{n-i-1} and S^{i+2}, joins (K_1^{i+1}) and (K_2^{i+1}).*[14]

---

[12] J. W. Alexander, *A proof and extension of the Jordan-Brouwer separation theorem*, Trans. Amer. Math. Soc., **23** (1922), pp. 333-349, Corollary W^i.

[13] I.e., a maximal connected subset of the complex representing the intersection.

[14] If K is a chain, (K) denotes the set of points in the chain. In connection with Lemma 1, see Lemma 4 of paper II. Since this paper was written I have noticed that Lemma 1 remains true if Γ^{n-i-1} denotes any closed point set. In this form, it can be shown to be *equivalent* to the Alexander Addition Theorem.

PROOF. As γ^i links Γ^{n-i-1}, Γ^{n-i-1} meets both K_1^{i+1} and K_2^{i+1}. Suppose no component of L^1 joins (K_1^{i+1}) and (K_2^{i+1}). Let us denote the sum of those components of L^1 that meet K_1^{i+1} by C_1, and the remainder by C_2. Also, let S^{i+2} be subdivided so finely that none of its closed cells meets both C_1 and C_2. The sum of those closed cells of S^{i+2} that meet C_1 form a chain $M^{i+2} \rightarrow F^{i+1}$. Then, from the relations

$$S^{i+2} \rightarrow K_1^{i+1} + K_2^{i+1}$$
$$M^{i+2} \rightarrow F^{i+1}$$

we get

$$S^{i+2} + M^{i+2} \rightarrow (K_1^{i+1} + F^{i+1}) + K_2^{i+1} \rightarrow 0 \qquad \text{(mod 2)}.$$

The last relation, together with the relation $K_2^{i+1} \rightarrow \gamma^i$, yields

$$K_1^{i+1} + F^{i+1} \rightarrow \gamma^i.$$

Now $S^{i+2} + M^{i+2}$ (mod 2) contains no points of C_1; indeed, its intersection with Γ^{n-i-1} contains only the points in C_2. No component of C_2 meets K_1^{i+1} by definition, and no component of C_2 meets F^{i+1} since the latter is a subchain of M^{i+2}. Consequently there are no points of L^1 in $K_1^{i+1} + F^{i+1}$. But this is equivalent to saying that γ^i does not link Γ^{n-i-1}, thus furnishing a contradiction.[15]

Suppose, now, that D is a domain and that C^i is an i-cell (possibly singular) whose vertices are $i + 1$ fixed points A_k of $\bar{D}$. By an *ϵ-removed realization of C^i in D* we mean a complex K^i constructed as follows: The vertices A_k are replaced by nearby (within a distance ϵ) points P_k in D; the 1-cells A_jA_k of C^i are replaced by 1-chains in D bounded by $P_j + P_k$ and lying in an ϵ-neighborhood of A_jA_k; and so on until we arrive at an $(i - 1)$-cycle γ^{i-1} in D. We make the convention that, during this process, a cell already in D is replaced by itself; thus, if $A_k \subset D$, then $P_k = A_k$; if $A_jA_k \subset D$, then the 1-chain replacing it is A_jA_k itself; and so on. Finally, K^i is an i-chain of D such that 1) $K^i \rightarrow \gamma^{i-1}$ and 2) each point on K^i is distant from some point of C^i by less than ϵ. The extension to ϵ-removed realizations of chains $M^i \rightarrow F^{i-1}$ such that if N^i and G^{i-1} are the replacements of M^i and F^{i-1} respectively, then $N^i \rightarrow G^{i-1}$, may then be carried out in such a way that each cell of N^i is an ϵ-removed realization in D of its correspond in M^i.

Concerning the *existence* of ϵ-removed realizations we prove the following basic lemma for cells; from it follows in an obvious manner the existence of ϵ-removed realizations of chains for domains of the type considered.

LEMMA 2. *Let D be a u.l.i-c. $(0 \leqq i \leqq n - 2)$ domain in E_n. Then for any*

---

[15] In proving this lemma for other moduli or for rational chains, it is of course necessary to make slight alterations; thus, the coefficients of the $(i + 2)$-cells in M^{i+2} must agree with the respective coefficients of the same cells in S^{i+2}, and we write $S^{i+2} - M^{i+2} \rightarrow (K_1^{i+1} - F^{i+1}) - K^{i+1} \rightarrow 0$, etc.

$\epsilon > 0$ there exists a $\delta > 0$ such that if C^j is a j-cell (possibly singular) $(0 \leqq j \leqq n-1)$ with vertices in $\bar{D}$ and which is of diameter less than δ, then there exists an ϵ-removed realization of C^j in D.

Proof. As the lemma is trivial for $j = 0$, we may use mathematical induction. We therefore assume the lemma true for cells of dimension $j - 1$.

Let $\delta' > 0$ be such that a $(j-1)$-cycle of D of diameter $< \delta'$ bounds a j-chain of D of diameter $< \epsilon/2$ (we assume $\delta' < \epsilon$). Also, let $\delta > 0$ be such that a $(j-1)$-cell with vertices on $\bar{D}$ and of diameter $< \delta$ has a $\delta'/4$-removed realization in D; we assume $\delta < \delta'$. Then if C^j is of diameter $< \delta/4$, each of its $(j-1)$-faces has a $\delta'/4$-removed realization in D, the sum of which yields a $(j-1)$-cycle γ^{j-1} of diameter $< \delta'$ constituting an ϵ-removed realization of the boundary of C^j. There exists a j-chain $K^j \to \gamma^{j-1}$ in D and whose diameter is $< \epsilon/2$. It is easily seen that K^j is an ϵ-removed realization of C^j in D.

It is interesting to note here, that while the assumption that D is u.l.i-c. for $0 \leqq i \leqq n-2$ is the important one from the point of view of our central problem, we have used, in the above proof, only the local connectedness for $0 \leqq i \leqq j-1$. Consequently,

Lemma 2a. *The conclusion of Lemma 2 is still true for a given j, under the (in general weaker) assumption that D is u.l.i-c. for $0 \leqq i \leqq j - j$.*

Theorem 2. *In E_n, let B be the common boundary of two domains D and D_1 each of which, for some fixed j $(0 \leqq j \leqq n-2)$ is u.l.i-c. for values of i such that $0 \leqq i \leqq j$. Then B is locally i-connected for the same values of i.*[16]

Proof. Let P be a point of B and ϵ an arbitrary positive number. By hypothesis there is a $\delta > 0$ such that any i-cycle $(0 \leqq i \leqq j)$ of D of diameter $\leqq \delta$ bounds an $(i+1)$-chain of D of diameter $< \epsilon$. I say that any cycle γ^i of $B \cdot S(P, \delta)$ bounds on $B \cdot \overline{S(P, \epsilon)}$. For suppose γ^i such a cycle non-bounding on $B \cdot \overline{S(P, \epsilon)}$. Then γ^i is linked with a cycle Γ^{n-i-1} of the complement of $B \cdot \overline{S(P, \epsilon)}$.[17] Thus, expressing γ^i as a Vietoris cycle $\{i_1, i_2, \cdots, i_k, \cdots\}$, we may assume Γ^{n-i-1} linked with each geometrically realized i_k. By Lemma 2a, we can choose η and k such that i_k has η-removed realizations γ_1^i and γ_2^i in $D \cdot S(P, \delta)$ and $D_1 \cdot S(P, \delta)$, respectively, and such that γ_1^i and γ_2^i together bound a chain H^{i+1} in $S(P, \delta)$ that does not meet Γ^{n-i-1}.

Let K_1^{i+1} and K_2^{i+1} be chains bounded by γ_1^i and γ_2^i in $D \cdot S(P, \epsilon)$ and $D_1 \cdot S(P, \epsilon)$, respectively. Also, let S^{i+2} be any chain bounded by $K_1^{i+1} + K_2^{i+1} + H^{i+1}$ in $S(P, \epsilon)$. Applying Lemma 1, some component, L^1, of the intersection of Γ^{n-i-1} and S^{i+2} joins (K_1^{i+1}) and $(K_2^{i+1} + H^{i+1})$. As Γ^{n-i-1} does not meet H^{i+1}, L^1 joins (K_1^{i+1}) and (K_2^{i+1}). But as $(K_1^{i+1}) \subset D$ and $(K_2^{i+1}) \subset D_1$,

[16] For the case $j = 0$ we have here a new proof, not dependent on properties of non-Jordanian continua, of the well-known theorem that a common boundary of two u.l.0-c. domains is a Jordan continuum. See R. L. Moore, Proc. Nat. Acad. Sci., **8** (1922), pp. 33–38; also R. L. Wilder, Trans. Amer. Math. Soc., **32** (1930), p. 644, II.

[17] See F. Frankl, *Topologische Beziehungen in sich kompakter Teilmengen euklidischen Räume zu ihren Komplementen* ··· , Wien. Akad. der Wiss., Math.-Naturw. Kl., Sitz., **136** (1927), pp. 689–699; and P. Alexandroff, loc. cit., pp. 166–168.

PROOF. As γ^i links Γ^{n-i-1}, Γ^{n-i-1} meets both K_1^{i+1} and K_2^{i+1}. Suppose no component of L^1 joins (K_1^{i+1}) and (K_2^{i+1}). Let us denote the sum of those components of L^1 that meet K_1^{i+1} by C_1, and the remainder by C_2. Also, let S^{i+2} be subdivided so finely that none of its closed cells meets both C_1 and C_2. The sum of those closed cells of S^{i+2} that meet C_1 form a chain $M^{i+2} \to F^{i+1}$. Then, from the relations

$$S^{i+2} \to K_1^{i+1} + K_2^{i+1}$$

$$M^{i+2} \to F^{i+1}$$

we get

$$S^{i+2} + M^{i+2} \to (K_1^{i+1} + F^{i+1}) + K_2^{i+1} \to 0 \qquad (\text{mod } 2)\,.$$

The last relation, together with the relation $K_2^{i+1} \to \gamma^i$, yields

$$K_1^{i+1} + F^{i+1} \to \gamma^i\,.$$

Now $S^{i+2} + M^{i+2}$ (mod 2) contains no points of C_1; indeed, its intersection with Γ^{n-i-1} contains only the points in C_2. No component of C_2 meets K_1^{i+1} by definition, and no component of C_2 meets F^{i+1} since the latter is a subchain of M^{i+2}. Consequently there are no points of L^1 in $K_1^{i+1} + F^{i+1}$. But this is equivalent to saying that γ^i does not link Γ^{n-i-1}, thus furnishing a contradiction.[15]

Suppose, now, that D is a domain and that C^i is an i-cell (possibly singular) whose vertices are $i + 1$ fixed points A_k of $\bar{D}$. By an *ϵ-removed realization of C^i in D* we mean a complex K^i constructed as follows: The vertices A_k are replaced by nearby (within a distance ϵ) points P_k in D; the 1-cells A_jA_k of C^i are replaced by 1-chains in D bounded by $P_j + P_k$ and lying in an ϵ-neighborhood of A_jA_k; and so on until we arrive at an $(i - 1)$-cycle γ^{i-1} in D. We make the convention that, during this process, a cell already in D is replaced by itself; thus, if $A_k \subset D$, then $P_k = A_k$; if $A_jA_k \subset D$, then the 1-chain replacing it is A_jA_k itself; and so on. Finally, K^i is an i-chain of D such that 1) $K^i \to \gamma^{i-1}$ and 2) each point on K^i is distant from some point of C^i by less than ϵ. The extension to ϵ-removed realizations of chains $M^i \to F^{i-1}$ such that if N^i and G^{i-1} are the replacements of M^i and F^{i-1} respectively, then $N^i \to G^{i-1}$, may then be carried out in such a way that each cell of N^i is an ϵ-removed realization in D of its correspond in M^i.

Concerning the *existence* of ϵ-removed realizations we prove the following basic lemma for cells; from it follows in an obvious manner the existence of ϵ-removed realizations of chains for domains of the type considered.

LEMMA 2. *Let D be a u.l.i-c. $(0 \leqq i \leqq n - 2)$ domain in E_n. Then for any*

[15] In proving this lemma for other moduli or for rational chains, it is of course necessary to make slight alterations; thus, the coefficients of the $(i + 2)$-cells in M^{i+2} must agree with the respective coefficients of the same cells in S^{i+2}, and we write $S^{i+2} - M^{i+2} \to (K_1^{i+1} - F^{i+1}) - K^{i+1} \to 0$, etc.

$\epsilon > 0$ *there exists a* $\delta > 0$ *such that if* C^j *is a j-cell (possibly singular)* $(0 \leqq j \leqq n-1)$ *with vertices in* $\bar{D}$ *and which is of diameter less than* δ, *then there exists an* ϵ*-removed realization of* C^j *in* D.

PROOF. As the lemma is trivial for $j = 0$, we may use mathematical induction. We therefore assume the lemma true for cells of dimension $j-1$.

Let $\delta' > 0$ be such that a $(j-1)$-cycle of D of diameter $< \delta'$ bounds a j-chain of D of diameter $< \epsilon/2$ (we assume $\delta' < \epsilon$). Also, let $\delta > 0$ be such that a $(j-1)$-cell with vertices on $\bar{D}$ and of diameter $< \delta$ has a $\delta'/4$-removed realization in D; we assume $\delta < \delta'$. Then if C^j is of diameter $< \delta/4$, each of its $(j-1)$-faces has a $\delta'/4$-removed realization in D, the sum of which yields a $(j-1)$-cycle γ^{j-1} of diameter $< \delta'$ constituting an ϵ-removed realization of the boundary of C^j. There exists a j-chain $K^j \to \gamma^{j-1}$ in D and whose diameter is $< \epsilon/2$. It is easily seen that K^j is an ϵ-removed realization of C^j in D.

It is interesting to note here, that while the assumption that D is u.l.i-c. for $0 \leqq i \leqq n-2$ is the important one from the point of view of our central problem, we have used, in the above proof, only the local connectedness for $0 \leqq i \leqq j-1$. Consequently,

LEMMA 2a. *The conclusion of Lemma 2 is still true for a given j, under the (in general weaker) assumption that D is u.l.i-c. for* $0 \leqq i \leqq j-j$.

THEOREM 2. *In* E_n, *let B be the common boundary of two domains D and* D_1 *each of which, for some fixed j* $(0 \leqq j \leqq n-2)$ *is u.l.i-c. for values of i such that* $0 \leqq i \leqq j$. *Then B is locally i-connected for the same values of i.*[16]

PROOF. Let P be a point of B and ϵ an arbitrary positive number. By hypothesis there is a $\delta > 0$ such that any i-cycle $(0 \leqq i \leqq j)$ of D of diameter $\leqq \delta$ bounds an $(i+1)$-chain of D of diameter $< \epsilon$. I say that any cycle γ^i of $B \cdot S(P, \delta)$ bounds on $B \cdot \overline{S(P, \epsilon)}$. For suppose γ^i such a cycle non-bounding on $B \cdot \overline{S(P, \epsilon)}$. Then γ^i is linked with a cycle Γ^{n-i-1} of the complement of $B \cdot \overline{S(P, \epsilon)}$.[17] Thus, expressing γ^i as a Vietoris cycle $\{i_1, i_2, \cdots, i_k, \cdots\}$, we may assume Γ^{n-i-1} linked with each geometrically realized i_k. By Lemma 2a, we can choose η and k such that i_k has η-removed realizations γ_1^i and γ_2^i in $D \cdot S(P, \delta)$ and $D_1 \cdot S(P, \delta)$, respectively, and such that γ_1^i and γ_2^i together bound a chain H^{i+1} in $S(P, \delta)$ that does not meet Γ^{n-i-1}.

Let K_1^{i+1} and K_2^{i+1} be chains bounded by γ_1^i and γ_2^i in $D \cdot S(P, \epsilon)$ and $D_1 \cdot S(P, \epsilon)$, respectively. Also, let S^{i+2} be any chain bounded by $K_1^{i+1} + K_2^{i+1} + H^{i+1}$ in $S(P, \epsilon)$. Applying Lemma 1, some component, L^1, of the intersection of Γ^{n-i-1} and S^{i+2} joins (K_1^{i+1}) and $(K_2^{i+1} + H^{i+1})$. As Γ^{n-i-1} does not meet H^{i+1}, L^1 joins (K_1^{i+1}) and (K_2^{i+1}). But as $(K_1^{i+1}) \subset D$ and $(K_2^{i+1}) \subset D_1$,

[16] For the case $j = 0$ we have here a new proof, not dependent on properties of non-Jordanian continua, of the well-known theorem that a common boundary of two u.l.0-c. domains is a Jordan continuum. See R. L. Moore, Proc. Nat. Acad. Sci., **8** (1922), pp. 33–38; also R. L. Wilder, Trans. Amer. Math. Soc., **32** (1930), p. 644, II.

[17] See F. Frankl, *Topologische Beziehungen in sich kompakter Teilmengen euklidischen Räume zu ihren Komplementen* ···, Wien. Akad. der Wiss., Math.-Naturw. Kl., Sitz., **136** (1927), pp. 689–699; and P. Alexandroff, loc. cit., pp. 166–168.

$L^1 \cdot B \cdot \overline{S(P, \epsilon)} \neq 0$. However, $L^1 \subset (\Gamma^{n-i-1}) \subset E_n - B \cdot \overline{S(P, \epsilon)}$, and the assumption that $\gamma^i \not\sim 0$ in $B \cdot \overline{S(P, \epsilon)}$ has led to a contradiction.

THEOREM 3. *In E_n, let B be the common boundary of two domains D and D_1, and j an integer $(0 \leqq j \leqq n - 2)$ such that both D and D_1 are u.l.i-c. for $0 \leqq i \leqq j$. Then if P is a point of B and ϵ an arbitrary positive number, there exist numbers δ and η, $\epsilon > \delta > \eta > 0$, such that 1) when $i < n - 2$ (as well as $\leqq j$), any cycle γ^i of $B \cdot F(P, \delta)$ bounds on $B \cdot [S(P, \epsilon) - S(P, \eta)]$; and 2) when $i = n - 2$ $(= j)$, any cycle γ^i of $B \cdot F(P, \delta)$ bounds on $B - B \cdot S(P, \eta)$.*

PROOF. We first consider the proof of 1). Let $\delta > 0$ be such that i-cycles of D and D_1 in $S(P, 2\delta)$ bound in $D \cdot S(P, \epsilon)$ and $D_1 \cdot S(P, \epsilon)$, respectively. Let γ^i be a cycle of $B \cdot F(P, \delta)$. Expressing γ^i as a Vietoris cycle, $\{i_1, i_2, \cdots, i_k, \cdots\}$, there exist $\alpha > 0$ and k such that i_k has α-removed realizations γ_1^i and γ_2^i in $D \cdot S(P, 2\delta)$ and $D_1 \cdot S(P, 2\delta)$, respectively, and such that all homologies $i_{k+m} \sim i_{k+m+1}$ $(m \geqq 0)$ are geometrically realizable in $D \cdot S(P, 2\delta)$ and $D_1 \cdot S(P, 2\delta)$. Let $\eta > 0$ be such that $\overline{S(P, \eta)}$ contains no points of K_1^{i+1}, K_2^{i+1}, (where $K_1^{i+1} \to \gamma_1^i$ in $D \cdot S(P, \epsilon)$, $K_2^{i+1} \to \gamma_2^i$ in $D \cdot S(P, \epsilon)$), or of any of the realizations of the homologies $i_{k+m} \sim i_{k+m+1}$ just referred to.

The cycle $\gamma^i \sim 0$ on $B[\overline{S(P, \epsilon)} - S(P, \eta)]$. For suppose not. Then it is linked with a cycle Γ^{n-i-1} of the *complement* of the latter set. There exist α and m such that i_k has α-removed realizations Γ_1^i and Γ_2^i not meeting Γ^{n-i-1} and such that $H^{i+1} \to \Gamma_1^i + \Gamma_2^i$ in $S(P, \epsilon) - S(P, \eta)$ does not meet Γ^{n-i-1}. Making use of the homologies connecting i_k with succeeding members of the cycle γ^i, we "extend" K_1^{i+1} and K_2^{i+1} so that they become new chains satisfying the same conditions as the old, except that $K_1^{i+1} \to \Gamma_1^i$ and $K_2^{i+1} \to \Gamma_2^i$.

Consider $S^{i+2} \to K_1^{i+1} + K_2^{i+1} + H^{i+1}$ in $S(P, \epsilon) - S(P, \eta)$. The intersection of S^{i+2} with Γ^{n-i-1} has, by Lemma 2, a component L^1 joining (K_1^{i+1}) to $(K_2^{i+1} + H^{i+1})$, which, since Γ^{n-i-1} fails to meet H^{i+1}, joins (K_1^{i+1}) to (K_2^{i+1}). As in the proof of Theorem 2, we see that, here, L^1 meets $B \cdot [\overline{S(P, \epsilon)} - S(P, \eta)]$. As we are thus led to a contradiction, we conclude that $\gamma^i \sim 0$ on $B \cdot [\overline{S(P, \epsilon)} - S(P, \eta)]$.

Thus, for any *fixed* γ^i on $B \cdot F(P, \delta)$ an $\eta > 0$ may be determined as above (with reference to the K^{i+1}'s) such that $\gamma^i \sim 0$ on $B \cdot [\overline{S(P, \epsilon)} - S(P, \eta)]$. Now let θ be the smaller of the numbers $\epsilon - \delta$, $\delta/2$. By Lemma 1 there exists a $\beta > 0$ such that any i-cell with vertices in $\bar{D}$ has a θ-removed realization in D. Let all cycles γ^i on $B \cdot F(P, \delta)$ be represented in the form $\{i_1, \cdots, i_k, \cdots\}$ such that i_1 is already a $\beta/3$-cycle. By a theorem of Vietoris,[6] only a finite number of these cycles are β-independent on $B \cdot F(P, \delta)$—denote Vietoris cycles to which they correspond by $\gamma_1^i, \cdots, \gamma_h^i$. Let $\eta(< \delta/2)$ be a positive number smaller than the η's determined as above with reference to the cycles $\gamma_1^i, \cdots, \gamma_h^i$.

If γ^i is a cycle of $B \cdot F(P, \delta)$, its i_1 is β-homologous to a linear combination of the i_1's corresponding to the cycles $\gamma_1^i, \cdots, \gamma_h^i$. More explicitly, if $\gamma_p^i = \{i_1^p, i_2^p, \cdots, i_k^p, \cdots\}$, there exists a β-complex ${}^{\beta}K^{i+1}$ such that

$${}^{\beta}K^{i+1} \to i_1 + c_1 i_1^1 + \cdots + c_p i_1^p + \cdots + c_h i_1^h, \text{ on } B \cdot F(P, \delta)$$

where the c's are either 1 or 0. We may now let K^{i+1} be a θ-removed realization of ${}^{\beta}K^{i+1}$ in D; K^{i+1} will also lie in $S(P, \epsilon) - S(P, \delta/2)$. It is now merely necessary to notice that by combining K^{i+1} with the $(i + 1)$-chains already determined for those γ_i^p's for which $c_p \neq 0$, we determine an $(i + 1)$-chain for γ^i which lies in $D \cdot [S(P, \epsilon) - S(P, \eta)]$. A similar $(i + 1)$-chain is of course obtained in D_1. Thus, starting with *any* γ^i on $B \cdot F(P, \delta)$, we may carry out the argument given previously to show that it bounds on $B \cdot [\overline{S(P, \epsilon)} - S(P, \eta)]$.

The proof of 2) offers no difficulty now. We proceed almost exactly as above, except that the chain $S^{i+2} = S^n \rightarrow K_1^{i+1} + K_2^{i+1} + H^{i+1} = K_1^{n-1} + K_2^{n-1} + H^{n-1}$ must be obtained in $E_n - \overline{S(P, \eta)}$, as $K_1^{n-1} + K_2^{n-1} + H^{n-1}$ might link $F(P, \epsilon) + F(P, \eta)$; and $\Gamma^{n-i-1} = \Gamma^1$ is a cycle of $E_n - [B - B \cdot S(P, \eta)]$.

It need hardly be pointed out that because of the way we chose δ and η, these constants may be so chosen as to serve uniformly for all values of i. As a matter of fact we might observe that, having found a δ satisfying Theorem 3, we can choose *any* δ' such that $\delta > \delta' > 0$ and a corresponding $\eta' > 0$ (dependent on δ') such that the Theorem is still true in terms of these new constants.

By virtue of Corollary 1 and Theorems 2 and 3, and the well-known fact[18] that a set M whose complement is just two domains of which it is the common boundary satisfies 1) of definition M^{n-1}, we have

PRINCIPAL THEOREM A. *In order that a point set M in E_n should be a generalized closed* $(n - 1)$*-manifold it is necessary and sufficient that its complement be the sum of two uniformly locally i-connected* $(0 \leqq i \leqq n - 2)$ *domains, one of which is bounded and of which M is the common boundary.*

§2

In this section we shall be concerned with the problem of showing that the Betti numbers of a generalized closed manifold boundary behave as one might expect. Throughout this section, B will denote a g.c. $(n - 1)$-m. in E_n, and D and D_1 its respective complementary domains.

THEOREM 4. *The Betti numbers $p^i(B)$, $0 \leqq i \leqq n - 2$, are all finite.*

PROOF. As a corollary to Lemma 2, there is a $\delta > 0$ such that any $(i + 1)$-cell of B has realizations in both D and D_1. Every cycle γ^i of B, $\gamma^i = \{i_1, i_2, \cdots, i_k, \cdots\}$, may be so chosen that the "ϵ-cycles" $i_1, i_2, \cdots, i_k, \cdots$ are all $\delta/3$-cycles. By the theorem of Vietoris quoted above, the cycles i_1 have a finite basis $i_1^1, \cdots, i_1^k$, with respect to δ-homologies in B. Let the cycle of which i_1^p is the first member be denoted by γ_p^i. Then if $\gamma^i = \{i_1, \cdots, i_k, \cdots\}$ is any i-cycle of B, $\gamma^i \sim c_1\gamma_1^1 + \cdots + c_k\gamma_1^k$ on B, for some choice of the c's $= 0, 1$.

For suppose not. We have, however, a δ-chain ${}^{\delta}K^{i+1}$ such that, for some choice of the c's

$${}^{\delta}K^{i+1} \rightarrow i_1 + c_1 i_1^1 + \cdots + c_k i_1^k \text{ on } B\,.$$

---

[18] See P. Alexandroff, loc. cit., for instance.

As $\gamma_1 + c_1\gamma_1^i + \cdots + c_k\gamma_k^i \not\sim 0$ on B, there is a cycle Γ^{n-i-1} in $E_n - B$ linked with it. Now, by taking realizations K_1^{i+1} and K_2^{i+1} of ${}^{\delta}K^{i+1}$ in D and D_1, respectively, (incidentally "extending" them through a number of the i_h^p's if necessary as indicated in the proof of Theorem 3 above), we may, by a procedure like that employed in the proofs of Theorems 2 and 3 (based on Lemma 1), show the existence of a connected subset L^1 of (Γ^{n-i-1}) joining (K_1^{i+1}) to (K_2^{i+1}), giving a contradiction as before. It follows that the numbers $p^i(B)$ are all finite.

Since, by the general duality theorem for closed sets,[19]

$$p^i(B) = p^{n-i-1}(E_n - B) = p^{n-i-1}(D) + p^{n-i-1}(D_2)\,, \qquad (0 \leqq i \leqq n-2)\,,$$

we have,

COROLLARY 2. *The Betti numbers $p^i(D)$ and $p^i(D_1)$, $0 \leqq i \leqq n-1$, are all finite.*

[For $i = 0, n-1$, this follows from the well-known duality for closed sets and the connectedness (Lennes-Hausdorff) of B.]

THEOREM 5. *The following duality relation holds:*

$$p^i(D) = p^{n-i-1}(D_1)\,, \qquad 1 \leqq i \leqq n-2\,.$$

PROOF. Since both domains D and D_1 are u.l.i-c., $0 \leqq i \leqq n-2$, and we use only this property in the proof, we need only show that $p^i(D) \geqq p^{n-i-1}(D_1)$.

Let $\gamma_1^{n-i-1}, \cdots, \gamma_k^{n-i-1}$ be linearly independent cycles of D_1. Let $S_1, S_2, \cdots, S_m, \cdots$, be a sequence of simplicial subdivisions of E_n, such that the maximum cell diameter of S_m converges to zero as m increases. The sum of those closed cells of S_m that meet B we denote by Π_m. We may assume that the maximum cell diameter of S_1 is already so small that Π_1 does not meet any of the γ's.

As the γ's are linearly independent in $E_n - \Pi_m$, there exists[20] in Π_m a set of cycles ${}^m\Gamma_1^i, \cdots, {}^m\Gamma_k^i$ with which the γ's are uniquely linked. In D, for m great enough (dependent on the u.l.i-c. of D), there is a set of cycles ${}^m\alpha_1^i, \cdots, {}^m\alpha_k^i$ that approximate the γ's, the degree of approximation being dependent on m but approaching zero as m increases. Furthermore, for m great enough, we shall have

$${}^m\alpha_h^i \sim {}^m\Gamma_h^i, (h = 1, 2, \cdots, k), \text{ in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}.$$

To see this, we may cover each point x of ${}^m\Gamma_1^i$, for instance, by a spherical neighborhood of radius $\frac{1}{2}\rho(x, \sum_{h=1}^k \gamma_h^{n-i-1})$. There will exist a number $\theta > 0$ such that for each x, $\frac{1}{2}\rho(x, \sum_{h=1}^k \gamma_h^{n-i-1}) \geqq \theta$. As implied in the process of approximation, the cells of the α's approach zero in diameter as m increases, so that, for m great enough, each cell C^i of ${}^m\Gamma_1^i$ as well as its correspond c^i of ${}^m\alpha_1^i$

[19] References may be found in footnote[28] of paper II; or see Lefschetz's *Topology* (Amer. Math. Soc. Coll. Pub., **12**).

[20] F. Frankl, loc. cit., and L. Pontrjagin, *Zum Alexanderschen Dualitätssatz*, Gött. Nach., (1927), pp. 315–322.

lies in some sphere S of the above type. Then we have the bounding relations for these cells,

$$C^i \to B^{i-1} \qquad \text{in } S$$

$$c^i \to b^{i-1} \qquad \text{in } S\,.$$

The cycles B^{i-1} and b^{i-1} together bound in S,

$$F^i \to B^{i-1} + b^{i-1} \qquad \text{in } S\,.$$

The cycle $C^i + c^i + F^i$ bounds in S;

$$V^{i+1} \to C^i + c^i + F^i \qquad \text{in } S\,; \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}\,.$$

Setting up these relations for each cell C^i and its correspond, and summing, we get

$$\sum V^{i+1} \to \sum C^i + \sum c^i + \sum F^i \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}\,.$$

But $\sum F^i = 0$, since the incidence between the chains V^{i+1} is the same as between the C^i's. Also, $\sum C^i = {}^m\Gamma_1^i$, etc. Consequently the last relation above becomes

$$\sum V^{i+1} \to {}^m\Gamma_1^i + {}^m\alpha_1^i \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}\,.$$

We recall, now, that for each h, γ_h^{n-i-1} is linked with ${}^m\Gamma_h^i$. Then ${}^m\alpha_h^i$ links γ_h^{n-i-1}, since if it did not, then, making use of the last relation above, ${}^m\Gamma_h^i$ would not link γ_h^{n-i-1}. For a similar reason, the α's are independent in

$$D \subset E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}\,.$$

Consequently

$$p^i(D) \geqq p^{n-i-1}(D_1)\,,$$

and the theorem is proved.

It will be noted that in the above proof we used the u.l.j-c. of D only for values of $j = 0, 1, \cdots, i-1$. An interesting by-product of the above proof is:

COROLLARY 3. *In E_n, let D be a domain and j a natural number such that D is u.l.i-c. for $0 \leqq i \leqq j-1$. Then $p^i(D) \geqq p^{n-i-1}(E_n - \bar{D})$ for $i = 1, 2, \cdots, j$.*

Furthermore, since a manifold in the classical sense is also a generalized closed manifold, we may state the following:

COROLLARY 4. *If M^{n-1} is a closed manifold in the classical sense, in E_n, and D and D_1 are its complementary domains, then $p^i(D) = p^{n-i-1}(D_1)$ $(1 \leqq i \leqq n-2)$.*

Continuing with our study of B, we may now state the following theorem:

As $\gamma_1 + c_1\gamma_1^i + \cdots + c_k\gamma_k^i \nsim 0$ on B, there is a cycle Γ^{n-i-1} in $E_n - B$ linked with it. Now, by taking realizations K_1^{i+1} and K_2^{i+1} of ${}^\delta K^{i+1}$ in D and D_1, respectively, (incidentally "extending" them through a number of the i_h^p's if necessary as indicated in the proof of Theorem 3 above), we may, by a procedure like that employed in the proofs of Theorems 2 and 3 (based on Lemma 1), show the existence of a connected subset L^1 of (Γ^{n-i-1}) joining (K_1^{i+1}) to (K_2^{i+1}), giving a contradiction as before. It follows that the numbers $p^i(B)$ are all finite.

Since, by the general duality theorem for closed sets,[19]

$$p^i(B) = p^{n-i-1}(E_n - B) = p^{n-i-1}(D) + p^{n-i-1}(D_2)\,, \qquad (0 \leqq i \leqq n-2)\,,$$

we have,

Corollary 2. *The Betti numbers $p^i(D)$ and $p^i(D_1)$, $0 \leqq i \leqq n-1$, are all finite.* [For $i = 0, n-1$, this follows from the well-known duality for closed sets and the connectedness (Lennes-Hausdorff) of B.]

Theorem 5. *The following duality relation holds:*

$$p^i(D) = p^{n-i-1}(D_1)\,, \qquad 1 \leqq i \leqq n-2\,.$$

Proof. Since both domains D and D_1 are u.l.i-c., $0 \leqq i \leqq n-2$, and we use only this property in the proof, we need only show that $p^i(D) \geqq p^{n-i-1}(D_1)$.

Let $\gamma_1^{n-i-1}, \cdots, \gamma_k^{n-i-1}$ be linearly independent cycles of D_1. Let $S_1, S_2, \cdots, S_m, \cdots$, be a sequence of simplicial subdivisions of E_n, such that the maximum cell diameter of S_m converges to zero as m increases. The sum of those closed cells of S_m that meet B we denote by Π_m. We may assume that the maximum cell diameter of S_1 is already so small that Π_1 does not meet any of the γ's.

As the γ's are linearly independent in $E_n - \Pi_m$, there exists[20] in Π_m a set of cycles ${}^m\Gamma_1^i, \cdots, {}^m\Gamma_k^i$ with which the γ's are uniquely linked. In D, for m great enough (dependent on the u.l.i-c. of D), there is a set of cycles ${}^m\alpha_1^i, \cdots, {}^m\alpha_k^i$ that approximate the γ's, the degree of approximation being dependent on m but approaching zero as m increases. Furthermore, for m great enough, we shall have

$${}^m\alpha_h^i \sim {}^m\Gamma_h^i\,, (h = 1, 2, \cdots, k),\text{ in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}\,.$$

To see this, we may cover each point x of ${}^m\Gamma_1^i$, for instance, by a spherical neighborhood of radius $\frac{1}{2}\rho(x, \sum_{h=1}^k \gamma_h^{n-i-1})$. There will exist a number $\theta > 0$ such that for each x, $\frac{1}{2}\rho(x, \sum_{h=1}^k \gamma_h^{n-i-1}) \geqq \theta$. As implied in the process of approximation, the cells of the α's approach zero in diameter as m increases, so that, for m great enough, each cell C^i of ${}^m\Gamma_1^i$ as well as its correspond c^i of ${}^m\alpha_1^i$

[19] References may be found in footnote[28] of paper II; or see Lefschetz's *Topology* (Amer. Math. Soc. Coll. Pub., **12**).

[20] F. Frankl, loc. cit., and L. Pontrjagin, *Zum Alexanderschen Dualitätssatz*, Gött. Nach., (1927), pp. 315–322.

lies in some sphere S of the above type. Then we have the bounding relations for these cells,

$$C^i \to B^{i-1} \qquad \text{in } S$$

$$c^i \to b^{i-1} \qquad \text{in } S.$$

The cycles B^{i-1} and b^{i-1} together bound in S,

$$F^i \to B^{i-1} + b^{i-1} \qquad \text{in } S.$$

The cycle $C^i + c^i + F^i$ bounds in S;

$$V^{i+1} \to C^i + c^i + F^i \qquad \text{in } S; \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}.$$

Setting up these relations for each cell C^i and its correspond, and summing, we get

$$\sum V^{i+1} \to \sum C^i + \sum c^i + \sum F^i \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}.$$

But $\sum F^i = 0$, since the incidence between the chains V^{i+1} is the same as between the C^i's. Also, $\sum C^i = {}^m\Gamma_1^i$, etc. Consequently the last relation above becomes

$$\sum V^{i+1} \to {}^m\Gamma_1^i + {}^m\alpha_1^i \qquad \text{in } E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}.$$

We recall, now, that for each h, γ_h^{n-i-1} is linked with ${}^m\Gamma_h^i$. Then ${}^m\alpha_h^i$ links γ_h^{n-i-1}, since if it did not, then, making use of the last relation above, ${}^m\Gamma_h^i$ would not link γ_h^{n-i-1}. For a similar reason, the α's are independent in

$$D \subset E_n - \sum_{h=1}^{k} \gamma_h^{n-i-1}.$$

Consequently

$$p^i(D) \geqq p^{n-i-1}(D_1),$$

and the theorem is proved.

It will be noted that in the above proof we used the u.l.j-c. of D only for values of $j = 0, 1, \cdots, i-1$. An interesting by-product of the above proof is:

COROLLARY 3. *In E_n, let D be a domain and j a natural number such that D is u.l.i-c. for $0 \leqq i \leqq j-1$. Then $p^i(D) \geqq p^{n-i-1}(E_n - \bar{D})$ for $i = 1, 2, \cdots, j$.* Furthermore, since a manifold in the classical sense is also a generalized closed manifold, we may state the following:

COROLLARY 4. *If M^{n-1} is a closed manifold in the classical sense, in E_n, and D and D_1 are its complementary domains, then $p^i(D) = p^{n-i-1}(D_1)$ $(1 \leqq i \leqq n-2)$.*

Continuing with our study of B, we may now state the following theorem:

THEOREM 6: *The Poincaré duality,*

$$p^i(B) = p^{n-i-1}(B) \qquad 1 \leqq i \leqq n-2$$

holds for the g.c. $(n-1)$-m. B.

PROOF. Considering a fixed i, we have from Theorem 5:

$$p^i(D) = p^{n-i-1}(D_1)$$
$$p^i(D_1) = p^{n-i-1}(D)\,.$$

Adding these relations, and remembering that $p^i(E_n - B) = p^i(D) + p^i(D_1)$, etc., we have

$$p^i(E_n - B) = p^{n-i-1}(E_n - B)\,.$$

By the duality theorem for closed sets and their complements,

$$p^i(B) = p^{n-i-1}(E_n - B)\,,$$

etc. The theorem follows at once.

Finally, we list in one theorem a number of properties of B that follow immediately from Theorem 5. Incidentally, it need hardly be remarked that these properties also hold for the case where we know B to be a closed $(n-1)$-manifold in the classical sense.

THEOREM 7. *If B is a g.c. $(n-1)$-m. in E_n, and D is either domain complementary to B, then*

1) *the number $p^i(D) + p^{n-i-1}(D)$ is a topological invariant of D;*

2) $\displaystyle\sum_{i=1}^{n-2} p^i(D) = \frac{1}{2}\sum_{i=1}^{n-2} p^i(B)$;

and if n is odd and D_1 is the other complementary domain,

3) $p^{\frac{n-1}{2}}(D) = p^{\frac{n-1}{2}}(D_1)$.

§3

We now return to the central problem, our aim being to characterize the manifold by means of the properties of *one* complementary domain.

THEOREM 8. *In E_n, let D be a uniformly locally i-connected domain $(0 \leqq i \leqq n-2)$ that is simply $(n-1)$-connected and whose boundary is bounded. Then the boundary, B, of D is a Jordan continuum.*[21]

PROOF. That B is connected follows easily from the fact that D is simply $(n-1)$-connected—indeed, this is the sole reason for the inclusion of this condition in our hypothesis (see Theorem 17 below).

[21] Jordan continuum = Peano continuum = continuous curve = locally 0-connected, compact continuum.

We have, then, to show that B is locally 0-connected. Suppose the contrary. Then the set $E_n - D$, which is obviously connected (being the sum of B and certain complementary domains whose boundaries are in B), is not locally 0-connected. For, suppose $E_n - D$ locally 0-connected. Since B is not locally 0-connected, there exist concentric $(n - 1)$-spheres K_1 and K_2, and a sequence of subcontinua of B, viz., B_ω, B_1, B_2, B_3, $\cdots$, such that 1) each of these continua lies in $K_1 + K_2 + I$ (where I is the "shell" domain bounded by K_1 and K_2) and contains at least one point of each of the spheres K_1, K_2; 2) no two of the continua B_i have a point in common, and, indeed, each is a component of $B \cdot (K_1 + K_2 + I)$; 3) B_ω is the sequential limiting set of the sequence $B_1, B_2, B_3, \cdots$.[22] Let P be a point of $B_\omega \cdot I$, and U a spherical neighborhood of P that lies wholly in I. Since D is uniformly locally 0-connected, and $E_n - D$ is locally 0-connected, there exist continua B_i and B_j $(i \neq j)$, which are joined *in* U by an arc t_1 through D and an arc t_2 through $E_n - D$, such that, indeed, t_1 and t_2 have the same endpoints a and b, $t_1 - a - b \subset D$, and $t_2 \subset E_n - D$. However, there exists in U a continuum (rel. to U) C which does not meet B and which separates a and b in U. This separating set, C, must lie wholly in D or wholly in $E_n - \bar{D}$. But if $C \subset D$ it cannot meet t_2, and if it lies in $E_n - \bar{D}$ it cannot meet t_1; either is a contradiction of the fact that C must meet both t_1 and t_2. Consequently the supposition that $E_n - D$ is locally 0-connected leads to a contradiction, and for B to fail to be locally 0-connected implies that $E_n - D$ fails to be locally 0-connected.

To continue: Since $E_n - D$ is not locally 0-connected, there exist concentric $(n - 1)$-spheres K_1 and K_2 and a sequence of subcontinua of $E_n - D$, viz., H_ω, H_1, H_2, $\cdots$, satisfying conditions (relative to $E_n - D$) analogous to those stated above for the B_i's. Let S denote an $(n - 1)$-sphere concentric with K_1 and K_2, with radius the arithmetic average of the radii of K_1 and K_2. Regarding S as an $(n - 1)$-complex, we make a sequence of simplicial subdivisions of it which we denote by S_1, S_2, S_3, $\cdots$, such that the mesh of S_i converges to zero as i increases.

Consider the 0-cells of S_i. They are finite in number, and consequently there exists a continuum, $H_{k(i)}$, of the sequence H_1, H_2, H_3, $\cdots$, which contains none of these 0-cells. Let C_i be any $(n - 1)$-simplex of S_i, and denote its vertices by $a_1, a_2, \cdots, a_n$. Consider a 1-cell $a_k a_j$ of the boundary of C_i; if it lies wholly in D or in $E_n - D$ we leave it undisturbed. If $a_k \subset D$ and $a_j \subset D$, but $a_k a_j \not\subset D$, we replace $a_k a_j$ by a 1-chain $K^1_{ij} \rightarrow a_k + a_j$ in D. If $a_k \subset E_n - D$ and $a_j \subset D$, we proceed as follows: Let a_k' be the last point on $a_k a_j$ in the order from a_k to a_j such that the segment $\overline{a_k a_k'} \subset E_n - D$ (in case $a_k = a_k'$ we have a degenerate segment). Let U be a small neighborhood of a_k' containing no point of $H_{k(i)}$, and a_k'' a point of D in U. By combining $\overline{a_k a_k'}$ with 1-chains bounded 1) by a_k' and a_k'' in U and 2) by a_k'' and a_j in D, we get the 1-chain $K^1_{ij} \rightarrow a_k + a_j$

[22] This property of continua that are not locally 0-connected was first noted, in the plane, by R. L. Moore, loc. cit. It is easily seen to hold for higher dimensions.

for this case. Finally, the case where both a_k and a_j are in $E_n - D$ but $a_ka_j \not\subset E_n - D$, is treated in an obvious manner. The upshot of this procedure is to replace a_ka_j by a 1-chain K^1_{ij} with the same boundary, that does not meet $H_{k(i)}$; and in general K^1_{ij} will be an approximation to a_ka_j in the sense that as i increases, the diameter of the point set $(a_ka_j) + (K^1_{ij})$ decreases uniformly over S_i (dependent on the uniform local 0-connectedness of D).

We next consider a 2-cell $a_ka_ja_h$ of the boundary of C_i. The 1-cells of its boundary have been replaced above by 1-chains forming a cycle

$$\Gamma^1 = K^1_{kj} + K^1_{kh} + K^1_{jh}.$$

Let L^2 be any 2-chain bounded by Γ^1 and approximating the 2-cell. The only case we have to consider is that where $(L^2)\cdot H_{k(i)} \neq 0$. Let

$$F = (\Gamma^1) + (L^2)\cdot(E_n - D),$$

and obtain a decomposition of F into mutually exclusive closed sets F_1 and F_2 such that $F_1 \supset (\Gamma^1)$ and $F_2 \supset (L^2)\cdot H_{k(i)}$. Denote the (positive) distance between the point sets F_1 and F_2 by η, and subdivide the cells of L^2 so that the resulting mesh is $<\eta/2$. The cells of this subdivision that meet F_1 give us a 2-chain, $M^2 \to \Gamma^1 + \beta^1$. The cycle β^1 does not meet F, and consequently is a cycle of D. Let N^2 be a 2-chain of D bounded by β^1. From the relations

$$M^2 \to \Gamma^1 + \beta^1,$$

$$N^2 \to \beta^1,$$

we get $K^2 = M^2 + N^2 \pmod 2 \to \Gamma^1$, and K^2 does not meet $H_{k(i)}$.

The procedure just outlined for 2-cells is now carried out for the 3-cells, 4-cells, etc., of the boundary of C_i, until, finally, we replace C_i itself in a like manner by an $(n-1)$-chain C'_i, this final step being dependent on the u.l. $(n-2)$-c. of D.

Carrying out the above process over all of S_i (each i-cell, $0 \leqq i \leqq n-1$, of S_i receiving only one replacement), we replace S_i by a cycle Γ^{n-1}_i which does not meet $H_{k(i)}$. Because of the u.l.i-c. of D, we can make the degree of approximation of Γ^{n-1}_i to S_i as small as we please by taking i great enough. Hence for i great enough, Γ^{n-1}_i separates K_1 and K_2.[23] But $H_{k(i)}$ contains at least one point of each of the spheres K_1, K_2, and accordingly meets Γ^{n-1}_i. Thus the supposition that B is not a Jordan continuum leads to a contradiction.

Digression to the case $n = 3$

At this point we may solve the central problem for the case $n = 3$ very easily, in that we can show that the boundary of a u.l.i-c. $(i = 0, 1)$ bounded (simply 2-connected) domain is a 2-dimensional manifold in the classical sense. For this purpose we prove the following theorem:

[23] S_i is linked with a cycle γ^0 consisting of a point of K_1 and a point of K_2. For i sufficiently great, $S_i \sim \Gamma^{n-1}_i$ in $E_n - (K_1 + K_2)$, and consequently Γ^{n-1}_i is linked with γ^0.

THEOREM 9a. *If the boundary, B, of a u.l.0-c. domain D in E_3 is a Jordan continuum, and D is simply 1-connected, then every simple closed curve of B separates B.*[24]

PROOF. Suppose J is a simple closed curve of B such that $B - J$ is connected. Let Γ^1 be an irreducible 1-cycle that links J, and let ϵ denote one-half the distance between J and (Γ^1). There exists[25] a connected open subset R of B which forms an ϵ-neighborhood (rel. to B) of J and such that $B - R$ is a Jordan continuum C. Clearly Γ^1 can meet B only in C.

There exists $\delta > 0$ such that all points of B within a distance δ of J lie in R. Let γ^1 be an irreducible 1-cycle of D that $\delta/2$-approximates J—such a cycle exists because of the u.l.0-c. of D. As D is simply 1-connected, there is a chain $M^2 \rightarrow \gamma^1$ in D. The cycles Γ^1 and γ^1 are linked, since both are irreducible and J is continuously deformable into γ^1 without meeting Γ^1.[26] Hence Γ^1 meets M^2 and there are points of Γ^1 in D. But Γ^1 is not wholly in D, else it could not link J (D being simply 1-connected). Then each point P of Γ^1 in D determines an open interval I of Γ^1 in D whose endpoints are in B. Only a finite number of distinct intervals I can meet M^2, since the distance between B and (M^2) is positive; denote such intervals by $I_1, I_2, \cdots, I_r$, the endpoints of $I_i (0 \leqq i \leqq r)$ by a_i, b_i.

As $a_i + b_i \subset C$, and C is a Jordan continuum, there is an arc t_i joining a_i and b_i in C. Because of the u.l.0-c. of D there exists (approximating t_i) in $D - (M^2)$ a 1-chain S_i with its boundary in I_i such that a) no point of S_i is within a distance δ of J; b) no two chains S_i meet and none meets Γ^i except in its own boundary; c) the portion of I_i that meets M^2 is entirely included in a 1-cycle Γ_i^1 formed from S_i and part of I_i. As D is simply 1-connected, $\Gamma_i^1 \sim 0$ in D, and therefore the chain $\Gamma^1 + \sum_{i=1}^r \Gamma_i^1$ links J. But then $\Gamma^1 + \sum_{i=1}^r \Gamma_i^1$ links γ^1, since no S_i enters a δ-neighborhood of J and consequently J is continuously deformable into γ^1 without meeting $\Gamma^1 + \sum_{i=1}^r \Gamma_i^1$. But this cycle fails to meet M^2 and we have a contradiction.

THEOREM 9b. *If B is the boundary of a u.l.i-c. ($i = 0, 1$) bounded domain D in E_3, there exists an $\epsilon > 0$ such that, P being any point of B, $\epsilon_1 \leqq \epsilon$, and C the component of $B \cdot S(P, \epsilon_1)$ determined by P, every simple closed curve of C separates C.*

PROOF. We choose ϵ so that any 1-cycle of D of diameter $< \epsilon$ bounds in D. Letting P and C be as stated above, suppose $C - J$ is connected. Let Γ^1

---

[24] Here, and in the case of the following theorem, we might generalize the argument to prove more general theorems. Thus, instead of 9a we might prove: *If the boundary, B, of a u.l.i-c. ($0 \leqq i \leqq n - 3$) domain D in E_n ($n > 2$) is a Jordan continuum, and D is simply* 1- *and* ($n - 2$)*-connected, then every closed cantorian* ($n - 2$)*-manifold* (Alexandroff, loc. cit.) *of B separates B.* I prefer, however, not to introduce generality where it seems pointless.

[25] R. L. Wilder, *On the imbedding of subsets of a metric space in Jordan continua*, Fund. Math., **19** (1932), pp. 45–64.

[26] See L. Pontrjagin, loc. cit.

link J in $S(P, \epsilon_1)$, and let R be chosen relative to C just as it was chosen relative to B in Theorem 9a. We also get γ^1 and $M^2 \to \gamma^1$ (in D) as in 9a, and arcs $t_i \subset C - R$ and the S_i's approximating the t_i's—all these, except M^2, being in $S(P, \epsilon_1)$ however. Then, by making use of the fact that Γ^1 also links J in E_n,[27] we can proceed to a contradiction as in 9a.

THEOREM 10a. *If a continuum B is the boundary of a u.l.i-c. ($i = 0, 1$) bounded domain D in E_3, no arc of B separates B.*

PROOF. Suppose t an arc of B such that $B - t$ is not connected. Then there exists a bounded continuum K which separates, in E_3, two points x and y of B, and such that $K \cdot (B - t) = 0$. Obviously $K \cdot B \subset t$. Then the 0-cycle $x + y$ links a Vietoris 2-cycle of K, viz., $\Gamma^2 = \{\gamma_1, \gamma_2, \cdots, \gamma_k, \cdots\}$. Since we may assume K is an *irreducible* cut between x and y, hence a common boundary of two domains, K is 2-dimensional at every point. Accordingly we may assume that for each k, the vertices of γ_k lie in $K - t$. Furthermore, $K - t$ is connected and therefore a subset of D (as is easily seen from the fact that it necessarily meets an arc of D joining points near x and y).

Let c_k be a 2-cell of γ_k, geometrically realized. The vertices of c_k are not in B; if a 1-cell of its boundary meets B, it may, for k great enough, be replaced by a small 1-chain of D with the same boundary. Similarly the intersection of c_k itself with B may be removed. In this way, we obtain an ϵ_k-modification, K_k^2, of γ_k, $\lim \epsilon_k = 0$ as $k \to \infty$, such that K_k^2 is in D. For k great enough, however, $x + y$ links K_k^2, and we have a contradiction.

THEOREM 10b. *Under the hypothesis of Theorem 10a, B is not locally separated by arcs; more explicitly, given a point P of B, and $\eta > 0$ arbitrary, no arc of C, the component of $B \cdot S(P, \eta)$ determined by P, separates C.*

PROOF. The proof resembles that of Theorem 10a with the chief exception that K is a continuum relative to $S(P, \eta)$ not meeting $C - t$ (although it may meet $B - C$).

For use below, we interpolate at this point the extension of Theorems 10a, 10b, which is obviously provable in the same manner:

THEOREM 10c. *In E_n, if a compact continuum B is the boundary of a u.l.i-c. domain ($0 \leqq i \leqq n - 2$), then B is not separated by any of its arcs, and not locally separated by its arcs.*

Since a Jordan continuum which is not disconnected by any of its arcs, but is disconnected by the omission of any of its simple closed curves is a simple closed surface,[28] we have, by combining Theorems 1, 8, 9a and 10a:

PRINCIPAL THEOREM B_1. *In order that a simply i-connected ($i = 1, 2$) bounded domain D in E_3 should have a simple closed surface (= topological 2-sphere) for its boundary, it is necessary and sufficient that D be uniformly locally i-connected for $i = 0, 1$.*

---

[27] See Lemma 2, p. 293, of paper II.

[28] L. Zippin, *On continuous curves and the Jordan Curve Theorem*, Amer. Jour. Math., 52 (1930), pp. 331–350.

Since a non-compact Jordan continuum which is not disconnected by any of its arcs but is disconnected by any of its simple closed curves is locally an E_2,[29] and a compact continuum which is locally an E_2 is a closed 2-dimensional manifold,[30] we have, combining Theorems 1, 8, 9b and 10b:

PRINCIPAL THEOREM B_2. *In order that a simply* 2*-connected bounded domain* D *in* E_3 *should have a two-dimensional closed manifold (in the classical sense) as its boundary, it is necessary and sufficient that* D *should be uniformly locally* i*-connected for* $i = 0, 1$.

Return to the general case

THEOREM 11. *Under the conditions stated in Theorem* 8, *every point of* B *is a limit point of* $E_n - \bar{D}$.

PROOF. Suppose P a point of B that is not a limit point of $E_n - \bar{D}$. Then there is an $\epsilon > 0$ such that $(E_n - \bar{D}) \cdot S(P, \epsilon) = 0$, $B \cdot F(P, \epsilon) \neq 0$, and all i-cycles $(0 \leqq i \leqq n - 2)$ of D of diameter $< 2\epsilon$ bound in D. Let $S = F(P, \epsilon/2)$. Let $S_1, S_2, \cdots, S_k, \cdots$ be a sequence of simplicial subdivisions of S, where the mesh of S_k is ϵ_k and $\lim \epsilon_k = 0$ as $k \to \infty$. Let $a_1, a_2, \cdots, a_n$ be the vertices of an $(n - 1)$-simplex of S_k. If a_h $(1 \leqq h \leqq n)$ is in D, we retain it and denote it by b_h; if a_h is in B, let b_h be a point of D within a distance η_k of a_h, where $\lim \eta_k = 0$ as $k \to \infty$. For k great enough, b_h and b_j

$$(h \neq j, 0 \leqq h \leqq n, 0 \leqq j \leqq n)$$

bound a (small) 1-chain of D which is to replace the 1-cell $a_h a_j$ in S_k. In the next step, the replaced boundaries of 2-cells bound, for k great enough, 2-chains of D which will replace the original 2-cells. This process is continued until the $(n - 1)$-cells of S_k are replaced by $(n - 1)$-chains of D, and the totality of these forms an $(n - 1)$-cycle Γ_k^{n-1} of D which δ_k-approximates S_k; and due to the *uniform* l.i-c. of D, $\lim \delta_k = 0$ as $k \to \infty$. Ultimately,[31] for k great enough, Γ_k^{n-1} must be linked with a point-pair consisting of P and a point of $B \cdot F(P, \epsilon)$. But this is impossible since Γ_k^{n-1} does not meet the continuum B. We must conclude, then, that no such point as P exists.

THEOREM 12. *Under the conditions stated in Theorem* 8, *the point set* $E_n - \bar{D}$ *is uniformly locally* 0*-connected.*

PROOF. Suppose $E_n - \bar{D}$ not u.l.0-c. Then it follows that there exists in B a point P, and an $\epsilon > 0$, such that for any $\delta > 0$ there is a point-pair in $(E_n - \bar{D}) \cdot S(P, \delta)$ that bounds no 1-chain in $(E_n - \bar{D}) \cdot S(P, \epsilon)$. Clearly we may suppose ϵ so small that all $(n - 2)$-cycles of D of diameter $< \epsilon$ bound in D. Let C denote the component of B determined by P in $S(P, \epsilon)$, and, since by Theorem 8, B is a Jordan continuum, let $\delta > 0$ be such that $B \cdot S(P, \delta) \subset C$. Let x and y denote distinct points of $(E_n - \bar{D}) \cdot S(P, \delta)$.

---

[29] L. Zippin, loc. cit.

[30] T. Radó, *Über den Begriff der Riemannsche Fläche*, Acta Litt. ac Scient. (Szeged), 2 (1925), pp. 101–121.

[31] See footnote[23].

Let xx' be a straight line interval from x to a point x' of C, in $S(P, \delta)$ and meeting B only at x'; let yy' be a similar interval from y to a point y' of C. Also, let t be an arc of C from x' to y'. By Theorem 10c, $B - t$ is connected. Let xx_1 and yy_1 be non-intersecting straight line intervals in $S(P, \delta)$, similar to xx' and yy' (but not meeting these), and meeting B only in the points x_1, y_1 of $B - t$. Let s be an arc of $B - t$ from x_1 to y_1. Let η be one-half the distance between t and s, and let $N(t)$ be an η'-neighborhood of t in $B \cdot S(P, \epsilon)$, where $\eta' < \eta$. Denote the boundary (rel. to B) of $N(t)$ by $F(t)$.

On t, in the order named, let $w_1, w_2, \cdots, w_k$ be a sequence of points such that $w_1 = x'$, $w_k = y'$, and $\delta(w_i w_{i+1}) < \frac{1}{2}\rho(t, B - N(t))$, where $w_i w_{i+1}$ is subarc of t. Let t_i denote a polygonal line approximating the straight line interval from w_i to w_{i+1}. On s, let $u_1, u_2, \cdots, u_h$ be a sequence such that $u_1 = x_1$, $u_h = y_1$, and $\delta(u_i u_{i+1}) < \frac{1}{2}\rho(S, \overline{N(t)} + \Sigma t_i)$. We join the u's by polygonal lines s_i. The above construction can be carried out so that, if K_1^1 and K_2^1 are chains defined as follows:

$$K_1^1 = xx' + yy' + \Sigma t_i,$$

$$K_2^1 = xx_1 + yy_1 + \Sigma s_i,$$

then $\Gamma^1 = K_1^1 + K_2^1$ is an irreducible 1-cycle. We note that

$$K_1^1 \to x + y \text{ in } E_n - [F(P, \epsilon) + B - N(t)]$$

$$K_2^1 \to x + y \text{ in } E_n - \overline{N(t)}\,.$$

Hence if Γ^1 does not link $[F(P, \epsilon) + B - N(t)] \cdot \overline{N(t)} = F(t)$, by the Alexander Addition Theorem $x + y \sim 0$ in $S(P, \epsilon) - B \cdot S(P, \epsilon)$. We shall show this to be the case.

Suppose Γ^1 links $F(t)$. Then it is linked with a Vietoris cycle γ^{n-2} of $F(t)$. As D is u.l.i-c., Γ^1 is linked with "ϵ-removed realizations" of γ^{n-2} in D. We choose one of these, Γ^{n-2}. As Γ^{n-2} is in $S(P, \epsilon)$, it bounds an M^{n-1} in D. As Γ^1 must meet M^{n-1}, we note that it meets it only in the t_i's or s_i's. However, such intersections are easily eliminated. For instance, suppose t_i meets M^{n-1}. There is a small spherical neighborhood S_i of t_i and the arc $w_i w_{i+1}$ of t, containing no point of Γ^{n-2}. We introduce new points of subdivision on $w_i w_{i+1}$, and new 1-chains joining these, that do not meet M^{n-1} and form a chain t_i'. As $t_i + t_i' \sim 0$ in S_i, hence in $E_n - \Gamma^{n-2}$, we may replace t_i in Γ^1 by t_i' without changing the linking property of Γ^1 and Γ^{n-2}. But after all such replacements (for both t_i's and s_i's) have been made, the new Γ^1 does not meet M^{n-1}, although it is still linked with Γ^{n-2}. This contradiction completes the proof.

THEOREM 13. *Under the conditions stated in Theorem* 8, $E_n - \bar{D}$ *is a connected point set.*

PROOF. By Theorem 8, B is a Jordan continuum. Let x and y be points of $E_n - \bar{D}$. On the straight line interval from x to y, let x_1 and y_1 be the first and last points of $\bar{D}$. The points x_1 and y_1 are points of B, and consequently are joined by an arc t of B. By Theorem 11, every point of t is a limit point

of $E_n - \bar{D}$, and by Theorem 12, $E_n - \bar{D}$ is u.l.0-c. In an obvious manner, then, we establish the existence, in $E_n - \bar{D}$, of an arc t' approximating t and joining a point x' of xx_1 to a point y' of yy_1. The portions xx' and yy' of xx_1 and yy_1, respectively, together with the arc t', constitute a continuum joining x and y in $E_n - \bar{D}$. Consequently $E_n - \bar{D}$ is connected.

We have now shown that under the conditions of Theorem 8, B is the common boundary of two domains D and $D_1 = E_n - \bar{D}$ whose sum is $E_n - B$, and hence $p^{n-1}(B) = 1$, whereas if F is a closed proper subset of B, $p^{n-1}(F) = 0$. Thus, *B satisfies condition* 1) *of definition M^{n-1}*.

By way of further comment on B, we note that by Theorem 4 of paper II, if F is a closed subset of B such that $p^{n-2}(F) = k$, then $B - F$ has at most $k + 1$ components; *in particular, if B contains an $(n - 2)$-sphere J^{n-2}, then $B - J^{n-2}$ has at most two components*, and *if B contains a closed $(n - 2)$-cell C^{n-2}, then $B - C^{n-2}$ is connected.* Consequently, J^{n-2} being as defined above, it follows easily from the facts just stated that J^{n-2} is the complete boundary of each component of $B - J^{n-2}$. From Theorem 10 and Corollary 3 of paper II, B has all these properties locally.

Not only is it true that the properties of B stated in the preceding paragraph are such as we would want for a manifold in the classical sense, but by generalizing Theorems 9a and 9b (see footnote[24]) we could show that under the generalized hypothesis of Theorem 9a, B satisfies the Jordan Curve Theorem (with respect to $(n - 2)$-spheres instead of simple closed curves, of course), and in any case it satisfies the Jordan Curve Theorem locally. In particular, if $n = 3$, we obtain in this manner another proof of Principal Theorems B_1 and B_2.

Returning to the central problem: In view of Principal Theorem A, in order to show that B, as defined in Theorem 8, is a g.c. $(n - 1)$-m., we need only establish the u.l.i-c. of D_1 for $1 \leqq i \leqq n - 2$. We now do this.

Theorem 14. *Under the hypothesis of Theorem* 8, *the domain $D_1 = E_n - \bar{D}$ is uniformly locally i-connected for* $1 \leqq i \leqq n - 2$.

Proof. Let ϵ be an arbitrary positive number; we may assume it small enough that cycles of D of diameter $< 2\epsilon$ bound in D. It is necessary only to show that if P is a point of B there exists a $\delta > 0$ such that any i-cycle of $D_1 \cdot S(P, \delta)$ bounds in $D_1 \cdot S(P, \epsilon)$. For brevity, let us denote the open set $D \cdot S(P, \epsilon)$ by Q, and let $\eta > 0$ be such that all cycles of $D \cdot S(P, 2\eta)$ bound in Q. Denote $F(P, \eta)$ by F. Let us select positive numbers satisfying the relation $0 < \epsilon_1 < \eta < \epsilon_2 < \epsilon$, and let N_1 denote $S(P, \epsilon_1)$.

Suppose there exists a cycle γ_1^i of $D_1 \cdot N_1$ that does not bound in $D_1 \cdot S(P, \epsilon)$. This cycle bounds in N_1; furthermore it bounds in $E_n - \bar{Q}$. For if γ_1^i did not bound in $E_n - \bar{Q}$ it would be linked with a cycle Γ^{n-i-1} of $\bar{Q}$, and consequently with an approximating cycle in $D \cdot S(P, 2\epsilon)$. As cycles of $D \cdot S(P, \epsilon)$ bound in D, hence in $E_n - (\gamma_1^i)$, this is impossible. We therefore have the relations

$$(1)\qquad \begin{array}{ll} K_{1\,1}^{i+1} \to \gamma_1^i \text{ in } N_1; & \text{in } E_n - \{F + [\bar{D} - \bar{D} \cdot S(P, \eta)]\} \\ K_{1\,2}^{i+1} \to \gamma_1^i \text{ in } E_n - \bar{Q}; & \text{in } E_n - \overline{D \cdot S(P, \eta)}\,. \end{array}$$

The cycle $\gamma_1^{i+1} = K_{11}^{i+1} + K_{12}^{i+1}$ links $\bar{D} \cdot F$, else, by the Alexander Addition Theorem, γ_1^i bounds in $S(P, \eta) - \bar{D} \cdot S(P, \eta)$ and hence in $D_1 \cdot S(P, \eta) \subset D_1 \cdot S(P, \epsilon)$. Therefore γ_1^{i+1} is linked with a cycle Γ_1^{n-i-2} of $\bar{D} \cdot F$.

Let α be a positive number $< \eta - \epsilon_1$ and $< \epsilon - \eta$. By Lemma 2 there is a $\delta > 0$ such that any cell of $\bar{D}$ of diameter $< \delta$ has an α-removed realization in D. Then in D there is a cycle β_1^{n-i-2} whose cells are of diameter $< \delta/3$, closely approximating Γ_1^{n-i-2}, and with which γ_1^{i+1} is linked. Now β_1^{n-i-2} bounds a chain M_1^{n-i-1} of Q. As this chain cannot meet K_{12}^{i+1}, it must meet K_{11}^{i+1}, and therefore $M_1^{n-i-1} \cdot N_1 \neq 0$.

Let $N_2 \subset S(P, \frac{1}{2}\epsilon_1)$ be a spherical neighborhood of P such that $N_2 \cdot (M_1^{n-i-1}) = 0$. Suppose that in N_2 there is a cycle γ_2^i of D_1 that fails to bound in $D_1 \cdot S(P, \epsilon)$. As above, we get chains K_{21}^{i+1} and K_{22}^{i+1} satisfying relations similar to (1) with respect to γ_2^i and N_2. The cycle $\gamma_2^{i+1} = K_{21}^{i+1} + K_{22}^{i+1}$ is linked with a Γ_2^{n-i-2} of $\bar{D} \cdot F$ as well as with a cycle β_2^{n-i-2} of D approximating Γ_2^{n-i-2} and similar to β_1^{n-i-2}. As above, we get $M_2^{n-i-1} \to \beta_2^{n-i-2}$ in Q; $(M_2^{n-i-1}) \cdot N_2 \neq 0$.

We continue as above; in the k^{th} step $(k > 1)$, N_k is an $S(P, \theta)$ with $\theta < \epsilon_1/k$ and such that $N_k \cdot (M_{k-1}^{n-i-1}) = 0$. We suppose γ_k^i in $D_1 \cdot N_k$ and failing to bound in $D_1 \cdot S(P, \epsilon)$. The chains K_{k1}^{i+1} and K_{k2}^{i+1}, forming γ_k^{i+1} are obtained, as well as the Γ_k^{n-i-2} of $\bar{D} \cdot F$, β_k^{n-i-2} and $M_k^{n-i-1} \to \beta_k^{n-i-2}$ in Q, with $(M_k^{n-i-2}) \cdot N_k \neq 0$, in a manner similar to that used in getting the corresponding chains above. We may suppose the cycles β_k^{n-i-2} all interior to a small spherical shell S, concentric with F, enclosing F and in $S(P, \epsilon_2) - S(P, \epsilon_1)$. If we let $D \cdot S = G$, then for all k, $\beta_k^{n-i-2} \subset G$.

By the theorem of Vietoris cited before, the cycles β_k^{n-i-2}, considered as $\delta/3$-cycles, are not δ-independent in $\bar{G}$. More precisely, there is a δ-complex ${}^{\delta}K^{n-i-1}$ bounded by some finite linear combination of the β's in $\bar{G}$. Thus,

$$\delta_K^{n-i-1} \to \beta_{k_1}^{n-i-2} + \beta_{k_2}^{n-i-2} + \cdots + \beta_{k_m}^{n-i-2} \text{ in } \bar{G},$$

where the β's are considered as $\delta/3$-cycles, and $k_1 < k_2 < \cdots < k_m$. In accordance with the way δ was selected, there is an α-removed realization of ${}^{\delta}K^{n-i-1}$ in D, where the realizations of the β's are the original (geometric) β's. That is,

$$K^{n-i-1} \to \beta_{k_1}^{n-i-2} + \beta_{k_2}^{n-i-2} + \cdots + \beta_{k_m}^{n-i-2} \text{ in } Q, \tag{2}$$

and K^{n-i-1} does not enter $S(P, \epsilon_1)$, and a fortiori does not enter $S(P, \epsilon_m)$.

From the relations

$$M_{k_p}^{n-i-1} \to \beta_{k_p}^{n-i-2}, \qquad p = 1, 2, \cdots, m-1,$$

combined with (2), mod 2, we get

$$K^{n-i-1} + \sum_{p=1}^{m-1} M_{k_p}^{n-i-1} \to \beta_{k_m}^{n-i-2} \text{ in } Q, \tag{3}$$

where the chain on the left of (3) does not enter $S(P, \epsilon_m)$. But then this chain does not meet $\gamma_{k_m}^{i+1}$, contradicting the fact that $\gamma_{k_m}^{i+1}$ and $\beta_{k_m}^{n-i-2}$ are linked with one another.

Consequently, we ultimately find a θ such that all i-cycles of $D_1 \cdot S(P, \theta)$ bound in $D_1 \cdot S(P, \epsilon)$, and D_1 is u.l.i-c.

As a result of Theorems 8, 11, 12, 13 and 14, combined with Principal Theorem A, we now have:

PRINCIPAL THEOREM C. *In E_n, in order that a bounded point set B should be a generalized closed $(n - 1)$-manifold, it is necessary and sufficient that it be the boundary of a uniformly locally i-connected* $(0 \leqq i \leqq n - 2)$ *domain that is simply $(n - 1)$-connected.*

Furthermore, from Principal Theorem C and Theorem 1, we have the following:

PRINCIPAL THEOREM D. *In E_n, let M be a bounded continuum satisfying conditions* 2) *and* 3) *of definition M^{n-1}; then each complementary domain of M has a generalized closed $(n - 1)$-manifold as its boundary.*

REMARKS. Principal Theorem D bears a close analogy to the Torhorst theorem[32] to the effect that if M is a Jordan continuum in E_2, then each complementary domain of M has a Jordan continuum as its boundary. For each g.c. $(n - 1)$-m. bounding a domain complementary to a set satisfying conditions 2) and 3) of definition M^{n-1}, in turn satisfies these conditions.

In particular, for the case $n = 3$, a set M as in Principal Theorem D has the property that its complementary domains are all bounded by closed 2-dimensional manifolds in the classical sense. One might say that such a configuration as M has an appearance similar to that of a piece of cheese, where the pores might be bounded by closed manifolds of various genuses.

In concluding this section we might also remark that for the plane case a quite simple proof of the theorem of Moore referred to in the introduction can be given using methods similar to those used above. In particular, the proof that the boundary of a u.l.0-c. bounded domain in the plane is a Jordan continuum is very simple (as in Theorem 8).

§4

In paper I, I suggested a possibility of generalizing the main result along lines similar to those taken in §1; i.e., by assuming a common boundary of two domains upon which we impose certain conditions. Principal Theorem C shows that the properties of one domain are sufficient for our generalization.

---

[32] Marie Torhorst, *Über den Rand der einfach zusammenhängenden ebenen Gebiete*, Math. Zeit., **9** (1921), pp. 44–65.

However, the question arises as to whether we can give generalizations of the main results of papers I, II and III which are more analogous to those results, by weakening the conditions on domain D of §3 and imposing conditions on domain D_1. This can be done in various ways as shown below.

LEMMA 3. *In E_n, let a bounded point set B be a common boundary of two uniformly locally 0-connected domains D and D_1, and let $p^{n-2}(E_n - B)$ be finite. Then both domains D and D_1 are uniformly locally $(n-2)$-connected.*

Lemma 3 is a direct consequence of Theorems 2a and 14 of paper II. By virtue of Lemma 3 and our preceding results[33] we have:

THEOREM 15. *In E_n, let a bounded point set B be a common boundary of (at least) two uniformly locally 0-connected domains D and D_1, such that $p^{n-2}(E_n - B)$ is finite and D is uniformly locally i-connected for $1 \leqq i \leqq n-3$; then B is a generalized closed $(n-1)$-manifold.*

By making use of Theorem 3, however, we can strengthen Theorem 15 so as to get the direct extension, to E_n, of Theorem 20 of paper II:

THEOREM 15a. *If, in Theorem 15, we replace the condition that $p^{n-2}(E_n - B)$ be finite, by the condition that $p^{n-2}(D)$ be finite, then the conclusion of that theorem still holds.*

PROOF. Let $\gamma_1^{n-2}, \gamma_2^{n-2}, \cdots, \gamma_k^{n-2}$ be a basis for the $(n-2)$-cycles of D. If P is a point of B, let $\epsilon > 0$ be such that no point of a $\gamma_h^{n-2} (1 \leqq h \leqq k)$ is on $S(P, \epsilon)$ or $F(P, \epsilon)$. With $j = 0$, let δ and η be as in Theorem 3. Also, let γ^{n-2} be a cycle of $D \cdot S(P, \eta)$. We have

$$K_1^{n-1} \rightarrow \gamma^{n-2} \text{ in } S(P, \eta)\text{ ; in } E_n - [F(P, \epsilon) + B - B \cdot S(P, \delta)]\,.$$

$$M_1^{n-1} \rightarrow \gamma^{n-2} + c_1\gamma_1^{n-2} + \cdots + c_k\gamma_k^{n-2} \text{ in } D\,.$$

$$M_2^{n-1} \rightarrow c_1\gamma_1^{n-2} + \cdots + c_k\gamma_k^{n-2} \text{ in } E_n - S(P, \epsilon)\,.$$

From the last two relations we obviously get a K_2^{n-1} such that

$$K_2^{n-1} \rightarrow \gamma^{n-2} \qquad\qquad \text{in } E_n - \overline{B \cdot S(P, \delta)}.$$

If the cycle $K_1^{n-1} + K_2^{n-1}$ does not link $B \cdot F(P, \delta)$, then γ^{n-2} bounds in $D \cdot S(P, \epsilon)$ by the Alexander Addition Theorem. And $K_1^{n-1} + K_2^{n-1}$ does not link $B \cdot F(P, \delta)$, since if it does it is linked with a 0-cycle of $B \cdot F(P, \delta)$, which is impossible since 0-cycles on $B \cdot F(P, \delta)$ bound in $B \cdot [S(P, \epsilon) - S(P, \eta)]$.

Thus D is u.l. $(n-2)$-c., and our result follows from Principal Theorem C.[33]

From Theorem 2a of paper II and Principal Theorem C we have the following extension of Theorem 8 of paper III.

THEOREM 16. *In E_n, let a bounded point set B be the common boundary of (at least) two uniformly locally 0-connected domains D and D_1, such that D is uniformly locally i-connected for $1 \leqq i \leqq n-3$ and $(n-2)$-cycles of D of diameter less than some ϵ bound in D. Then B is a generalized closed $(n-1)$-manifold.*

---

[33] That D is simply $(n-1)$-connected follows from the fact that a common boundary of two domains is connected.

Now the hypotheses of Theorems 15–16 suggest, certainly, that by further strengthening the conditions on D_1 we may weaken still more the conditions on D. So, in concluding this section, we give the following general formula for the complementary conditions on D and D_1:

Principal Theorem E. *In $E_n (n > 2)$, let a bounded point set B be a common boundary of (at least) two domains D and D_1, and let j be a fixed number $(n - 2 > j \geqq (n - 3)/2)$ such that* 1) *D is uniformly locally i-connected for $0 \leqq i \leqq j$, and the Betti numbers $p^{j+1}(D)$, $p^{j+2}(D)$, $\cdots$, $p^{n-2}(D)$ are all finite, and* 2) *D_1 is uniformly locally i-connected for $0 \leqq i \leqq n - j - 3$. Then B is a generalized closed $(n - 1)$-manifold.*

The proof is based on Principal Theorem C, and consists in showing that D is u.l.i-c. for $j + 1 \leqq i \leqq n - 2$. The latter is done just as for the special case $i = n - 2$ in the proof of Theorem 15a, on the basis of Theorem 3.

§5

We devote this section to a consideration of the most general open subsets of E_n that satisfy certain properties of local connectedness. In brief, we shall obtain generalizations to the case of general open sets, of Theorem 5 and Principal Theorem C.

Let D denote a bounded open subset of E_n that is u.l.i-c. for $0 \leqq i \leqq n - 2$. We note first that D has only a finite number of components. For suppose $P_1, P_2, P_3, \cdots$ were an infinite sequence of points from distinct components of D. As D is bounded, there would be a limit point, P, of the set $\sum_{i=1}^{\infty} P_i$. Let ϵ be any positive number and $\delta > 0$ such that 0-cycles of D of diameter $< \delta$ bound chains of D of diameter $< \epsilon$. Then for P_i and P_j $(i \neq j)$ in $S(P, \delta)$, there is an arc from P_i to P_j in D, contradicting the supposition that P_i and P_j were in distinct components of D.

Denoting the boundary of D by B, suppose B has infinitely many components. Then, as we shall show, $E_n - D$ has infinitely many components. For suppose $E_n - D$ has only a finite number, k, of components; denote them by $C_1, C_2, \cdots, C_k$. Then, since $E_n - D \supset B$, C_1, say, must contain infinitely many components of B; let B_1 and B_2 denote two such components. There is a bounded continuum, L, separating B_1 and B_2 such that $L \cdot B = 0$. As $L \cdot C_1 \neq 0$, $L \subset E_n - \bar{D}$. Now in arbitrarily small neighborhoods of B_1 and B_2, respectively, there are points of D which must belong to different components of D since D cannot meet L. We have shown, above, that D has only a finite number of components, say $D_1, D_2, \cdots, D_m$. Arbitrarily near to B_1 there are points of, say, D_1, and as already noted, D_1 has no limit points in B_2. Thus, we see that at most m components of B in C_1 contain limit points of D, which is absurd since every point of B is a limit point of D. Consequently, if B has infinitely many components, so, too, has $E_n - D$.

Conversely, if $E_n - D$ has infinitely many components, so, too, has B. Let

C be a component of $E_n - D$, and let ϵ be an arbitrary positive number. Also, let $N(C, \epsilon)$ denote an "ϵ-neighborhood" of C, which we may assume connected since C is connected. Let P be a point of $N(C, \epsilon)$ not in C, and PQ an arbitrary arc of $N(C, \epsilon)$ having only Q in common with C. The arc PQ is not wholly in $E_n - D$, else $P \subset C$. Consequently there exists a point P_1 of PQ such that $P_1 \subset D$. Hence the arc P_1Q, subset of PQ, contains a point of B. Thus, arbitrarily near to C there are points of B, and as B is closed, C contains at least one point of B. As $E_3 - D \supset B$, a component of B is a subset of one component of $E_3 - D$, and as each component of $E_3 - D$ contains at least one component of B, B has infinitely many components.

EXAMPLE. That there is, to be sure, a certain amount of arbitrariness in the number of components of D, B and $E_3 - D$, respectively, the following example shows: In the cartesian plane, let D consist of three components D_1, D_2, D_3 as follows: 1) D_1 is the set of points (x, y) such that $x^2 + y^2 < 1$; 2) D_2 is the set of points (x, y) such that $(x - 3)^2 + y^2 < 1$; 3) D_3 is the set of points (x, y) such that $36 > x^2 + y^2 > 25$. Here, B has four components, $E_2 - D$ has two components, and, as already noted, D has three components.

To continue with our investigation, let B_1 be a component of B, and let P be a point of B_1. As P is a limit point of D, and D has only a finite number of components, P is a limit point of one of these components, say of D_1. Then every point of B_1 is on the boundary of D_1. For let Q be any other point of B_1. Choosing an arbitrary $\epsilon > 0$, let $\delta > 0$ $(\epsilon > \delta)$ be such that any two points of D whose distance apart is $<\delta$ are joined by an arc of D of diameter $<\epsilon$. As B_1 is connected, there exist in B_1 points $P_1, P_2, \cdots, P_m$, where $P_1 = P$ and $P_m = Q$, such that $\rho(P_i, P_{i+1}) < \delta/4$, $(1 \leqq i \leqq m - 1)$. For each i, let P'_i be a point of D such that $\rho(P_i, P'_i) < \delta/4$, P'_1 being in D_1, and, noting that necessarily $\rho(P'_i, P'_{i+1}) < \delta$, let t_i be an arc of D from P'_i to P'_{i+1}. The point set $N = \sum_{i=1}^{m-1} t_i$ is a subset of D, and being connected lies entirely in one component of D, that component necessarily being D_1, which contains P'_1. As ϵ was arbitrary, we have shown that in every $S(Q, \epsilon)$ there is a point (P'_m) of D_1, so that Q is a boundary point of D_1. Finally, no other component of D has a boundary point in B_1, since, because of the u.l.0-c. of D, no two components have a common boundary point.

We summarize what we have shown in this section thus far in Lemma form, noting that we have used only the u.l.0-c. of D:

LEMMA 4. *Let D be a u.l.0-c., bounded, open subset of E_n. Then* 1) *D has only a finite number of components,* 2) *the boundary, B, of D has a finite or infinite number of components according as the number of components of $E_n - D$ is finite or infinite, and* 3) *if B_1 is a component of B, then B_1 is completely on the boundary of one and only one component of D.*

In view of the above, we may reduce the study of the general case to the case where D is connected. For, firstly, in discussing a component of the boundary of D_1, say, we are at the same time discussing a component of the boundary of D. Secondly, D_1 is u.l.i-c., as follows easily from the fact that a boundary point

of D_1 is not on the boundary of any other component of D. Consequently *we may assume from now on that D is a (connected) domain and B its boundary.*

Our next step is to show that if C is a component (domain) of $E_n - \bar{D}$, then the complete boundary of C is in one component of B. As the boundary of C is in B, let B_1 be a component of B containing a boundary of point of C, and suppose B_2 is any other component of B. There exists a continuum K separating B_1 and B_2 and such that $K \cdot B = 0$. As D has limit points in both B_1 and B_2, $D \cdot K \neq 0$ and consequently $D \supset K$. Then $C \subset E_n - K$, and therefore C cannot have a limit point in B_2. Thus B_1 contains the complete boundary of C.

Hereafter, if B_i is a component of B, we shall denote by B_i^* the point set consisting of B_i and all components of $E_n - \bar{D}$ whose boundaries lie in B_i. Such a set B_i^* is clearly a continuum, and only one, which we denote by B_u^*, is unbounded. The set $B^* = \sum_B B_i^* = E_n - D$. Also, a set B_i^* is a component of $E_n - D$, since any two sets $B_i, B_j (i \neq j)$ are separated by a continuum $K \subset D$ (as in the preceding paragraph) and consequently B_i^*, B_j^* are separated by K.

We now prove that if $\epsilon > 0$ is arbitrarily given, then only a finite number of components of B^* are of diameter $> \epsilon$. The components of B^* are the sets B_i^*. Suppose there exist infinitely many sets, $B_1^*, B_2^*, B_3^*, \cdots$, all of diameter $> \epsilon$; we may assume B_u^* not in this sequence, so that the sum $\sum_{i=1}^{\infty} B_i^*$ is bounded. For each i, let P_i and Q_i denote points of B_i^* such that $\rho(P_i, Q_i) > \epsilon$. It follows easily that there exist two points P and Q such that $\rho(P, Q) \geqq \epsilon$, and a subsequence $B_{k_i}^*$ $(i = 1, 2, 3, \cdots)$ of the above sequence such that P and Q are, respectively, the sequential limit points of the sequences $\{P_{k_i}\}$ and $\{Q_{k_i}\}$. Let K_1, K_2 and S denote the spheres $S(P, \epsilon/2)$, $S(P, \epsilon/4)$ and $S(P, 3\epsilon/8)$, respectively. We may now establish a contradiction just as in the proof of Theorem 8, where the continua B_i of that proof become, in the present instance, subcontinua of the sets $B_{k_i}^*$.

Now let B_1 be an arbitrary component of B, and denote by D_1 that domain complementary to B_1 that contains D. We first note that $D_1 \cdot B_1^* = 0$. For suppose P is a point of $D_1 \cdot B_1^*$. As D_1 contains no points of B_1, $P \subset B_1^* - B_1$, and by definition of B_1^*, P is therefore a point of some component, C, of $E_n - \bar{D}$, whose complete boundary is in B_1. But $D_1 \supset D$, and if Q is a point of D there is an arc of D_1 from P to Q which does not meet B_1. But this is absurd, since $P \subset C$ and $Q \subset E_n - C$. Thus $D_1 \cdot B_1^* = 0$. Moreover, $D_1 + B_1^* = E_n$. For if C is a component of $E_n - \bar{D}$ whose boundary is in B_1, then $B_1^* \supset C$; if C is a component of $E_n - \bar{D}$ whose boundary is in $B_2 \neq B_1$, then both C and B_2 lie in D_1; and all components of B distinct from B_1 lie in D_1, whereas $D \subset D_1$. Finally, we note that since B_1 is on the boundary of D, it is the complete boundary of D_1.

The domain D_1 is u.l.i-c. for $0 \leqq i \leqq n - 2$. For let $\epsilon > 0$ be arbitrary, and $\delta > 0$ such that i-cycles of D of diameter $< \delta$ bound chains of D of diameter $< \epsilon$. Let P be a point of B_1, and γ^i an irreducible cycle of $D_1 \cdot S(P, \delta)$. Also, let K^{i+1} be a chain of $S(P, \delta)$ bounded by γ^i and irreducible with respect to its

boundary (see paper II, p. 298). The point set (K^{i+1}) is connected (paper II, Lemma 5). If γ^i is a cycle of D, it bounds in $D \cdot S(P, \epsilon)$; we suppose, therefore, that γ^i meets $D_1 - D$ (hence also $B - B_1$). We assume that γ^i does not meet B_u^*; for if $B_1^* = B_u^*$, γ^i certainly does not meet it, and if $B_1^* \neq B_u^*$, $B_1^* \cdot B_u^* = 0$ as shown above, and we may assume δ so small as to exclude all points of B_u^*.

Consider the closed point set $F = (K^{i+1}) \cdot (E_n - D) + (\gamma^i)$. Let C denote the component of F determined by (γ^i); $C \cdot B_1 = 0$, else a connected subset of $E_n - D$ joins a point of (γ^i) to B_1, thus placing a point of (γ^i) in B_1^* (since, as shown above, B_1^* is a component of $E_n - D$), contrary to the fact that $D_1 \cdot B_1^* = 0$. Let $F = F_1 + F_2$ be a decomposition of F into mutually exclusive closed sets such that $F_1 \supset B_1^* \cdot (K^{i+1})$ and $F_2 \supset C$. Let η denote the distance $\rho(F_1, F_2)$, and make a subdivision of K^{i+1} into cells of mesh $< \eta/2$. From the cells of this subdivision that meet F_2 we form a chain M^{i+1} whose boundary is $\gamma^i + \Gamma^i$, where $\Gamma^i \subset D \cdot S(P, \delta)$. We have

$$M^{i+1} \to \gamma^i + \Gamma^i \qquad \text{in } (E_n - B_1^*) \cdot S(P, \epsilon)\,,$$
$$N^{i+1} \to \Gamma^i \qquad \text{in } D \cdot S(P, \epsilon)\,.$$

As $B_1^* + D_1 = E_n$, we get, mod 2,

$$M^{i+1} + N^{i+1} \to \gamma^i \qquad \text{in } D_1 \cdot S(P, \epsilon)\,.$$

Thus, as B_1 is the boundary of D_1, it follows from Principal Theorem C that *B_1 is a g.c. $(n-1)$-m.*

The structure of B_1^* is now seen to be rather simple. For B_1, being a g.c. $(n-1)$-m., has just two complementary domains, one of which is the D_1 defined above; the other, which we denote by D_2, is the only component of $E_n - \bar{D}$ which is added to B_1 to make up B_1^*. That is, $B_1^* = B_1 + D_2$. We can now show that, except for a finite number of them, the components B_k of B satisfy the relations $p^i(B_k) = 0$ for $1 \leqq i \leqq n-2$. There exists an $\epsilon > 0$ such that all i-cycles of D of diameter $< \epsilon$ bound in D. As has been shown above, only a finite number of components of B have a diameter $\geqq \epsilon$. Let B_1 be a component of B of diameter $< \epsilon$; we assume B_1 not a single point, else there is no problem. We define D_1 and D_2 as above, where $D_1 \supset D$. Now suppose B_1 carries a cycle γ^i that fails to bound on B_1. Accordingly there is a cycle γ^{n-i-1} in $E_n - B_1$ that fails to bound in $E_n - B_1$; we may assume γ^{n-i-1} wholly in either D_1 or D_2, since $1 \leqq n - i - 1 \leqq n - 2$. In either case, by employing the methods of approximation as in the proof of Theorem 5, we show the existence, in D, of a cycle γ^j ($j = i$ or $n - i - 1$) that fails to bound in D_1 and is of diameter $< \epsilon$. But if γ^j fails to bound in D_1 it certainly fails to bound in D, thus contradicting the fact that i-cycles of diameter $< \epsilon$ bound in D. Summarizing our results, we have the following:

Theorem 17. *Let D be a bounded, uniformly locally i-connected* $(0 \leqq i \leqq n-2)$, *open subset of E_n. Then* 1) *D consists of a finite number of domains $D_1, D_2, \cdots, D_m$ such that $\bar{D}_j \cdot \bar{D}_k = 0$ if $j \neq k$;* 2) *each component of B_k, the boundary of D_k* $(1 \leqq k \leqq m)$ *is either a point or a generalized closed* $(n-1)$*-manifold, and,*

furthermore, $B_k = B_{k0} + B_{k1} + \cdots + B_{kh} + B_{k(h+1)} + \cdots$, *where* B_{k0} *is the set of point-components of* B_k, *and the manifolds* B_{kj}, *where* $j > h$, *satisfy the conditions* $p^i(B_{kj}) = 0$ *for* $1 \leqq i \leqq n - 2$; 3) *the diameters of the manifolds* B_{kj} *converge to zero as* j *increases.*

The following corollary to Theorem 17 is worthy of note:

COROLLARY 4. *For the case* $n = 3$, *under the hypothesis of Theorem* 17, *the sets* B_{kj}, *for* $j > 0$, *are all closed two-dimensional manifolds in the classical sense, whereas those for which* $j > h$ *are true topological* 2*-spheres.*

We turn now to the converse of Theorem 17. That is, assuming the conditions stated in 1), 2) and 3) of the conclusion of that theorem, we shall prove that D is u.l.i-c. for $0 \leqq i \leqq n - 2$. Since $\bar{D}_j \cdot \bar{D}_k = 0$ if $j \neq k$, we may confine our attention to one component of $\bar{D}$, and, to simplify notation, we suppose that D itself is connected. Let B_1 denote, then, a component of B, and P a point of B_1, and let $\epsilon > 0$ be arbitrary. Because of condition 3), there exists a $\delta > 0$ such that if N is a component of B, distinct from B_1, which enters $\overline{S(P, \delta)}$, then $N \subset S(P, \epsilon)$. Also, as B_1 is either a point or a g.c. $(n - 1)$-m., there exists an $\eta_1 > 0$ such that if γ^i $(0 \leqq i \leqq n - 2)$ is a cycle in $S(P, \eta_1) - B_1 \cdot S(P, \eta_1)$, then $\gamma^i \sim 0$ in $S(P, \delta) - B_1 \cdot S(P, \delta)$ (Theorem 1).

There exists an $\eta_2 > 0$ such that if γ^i is a cycle of $D \cdot S(P, \eta_2)$, then $\gamma^i \sim 0$ in D. To see this, consider a γ^i that fails to bound in D and therefore links B. Then γ^i links a component of B.[34] Also $i > 0$, since all 0-cycles bound in D. By condition 2), γ^i must link one of the sets $B_{k1}, \cdots, B_{kh}$. But each B_{kj} $(1 \leqq j \leqq h)$ is a g.c. $(n - 1)$-m., and there exists a number $\alpha_j > 0$ such that no cycle γ^i for $i > 0$ of diameter $< \alpha_j$ links B_{kj}. Consequently if η_2 is the greatest lower bound of the numbers α_j $(1 \leqq j \leqq h)$, no γ^i of diameter $< \eta_2$ links B.

Let η denote the smaller of the numbers η_1, η_2. Then if γ^i $(0 \leqq i \leqq n - 2)$ is a cycle of $D \cdot S(P, \eta)$, $\gamma^i \sim 0$ in $D \cdot S(P, \epsilon)$. In the first place, there exists a $K_1^{i+1} \to \gamma^i$ in $S(P, \delta) - B_1 \cdot S(P, \delta)$. The set of points $T = B \cdot (K_1^{i+1})$ is closed; furthermore, if x is a point of T and N_x the component of B determined by x, the set of points $R = \sum_{x \subset T} N_x$ is closed by virtue of condition 2), and a subset of $S(P, \epsilon)$ as arranged for above. Clearly $R \subset B - B_1$. There exists a decomposition of B into mutually exclusive closed sets F_1 and F_2 such that $F_1 \supset R$, $F_1 \subset S(P, \epsilon)$, and $F_2 \supset B_1$. We have, then,

$$K_1^{i+1} \to \gamma^i \qquad \text{in } E_n - [F(P, \epsilon) + F_2].$$

Also, since $\eta \leqq \eta_2$, there is a chain K_2^{i+1} such that

$$K_2^{i+1} \to \gamma^i \qquad \text{in } D; \qquad \text{in } E_n - F_1.$$

As $[F(P, \epsilon) + F_2] \cdot F_1 = 0$, the cycle $K_1^{i+1} + K_2^{i+1}$ cannot link this product, and

[34] In general, if F is a bounded closed subset of E_n, and γ^i a cycle linking F, then γ^i links some component of F. For by an induction theorem of Brouwer (Proc. Amst. Acad., 14, pp. 137–147), F has a closed subset F_1 that is irreducibly linked by γ^i. That F is connected follows easily from the Alexander Addition Theorem.

by the Alexander Addition Theorem it follows that $\gamma^i \sim 0$ in $E_n - [F(P, \epsilon) + B]$, i.e., in $D \cdot S(P, \epsilon)$. Thus, D is u.l.i-c., and we have

CONVERSE OF THEOREM 17. *Under the conditions* 1), 2) *and* 3) *of the conclusion of Theorem* 17, *the open set D is uniformly locally i-connected* $(0 \leqq i \leqq n - 2)$.

Theorem 17 and its converse constitute a generalization of Principal Theorem C; in the latter case, the added assumptions of the connectedness and simple $(n - 1)$-connectedness of D limit B to one g.c. $(n - 1)$-m.

If to the hypothesis of Theorem 17 *we add the condition that D is u.l. $(n - 1)$-connected, the conclusion is modified as follows:* 1) *remains as before;* 2) *becomes "each component of B_k, the boundary of B_k, is a g.c. $(n - 1)$-m., and there are only a finite number of such components;"* 3) *becomes redundant.* The proof is simple: If there were infinitely many components B_k, or a point component of B, there would be, as shown above, a B_k of diameter $< \epsilon$, where ϵ is such that $(n - 1)$-cycles of D of diameter $< \epsilon$ bound in D. Using separation properties of closed sets as we have done before, we easily establish the existence, in D, of a cycle γ^{n-1} approximating B_k and linking a 0-cycle of B; the contradiction arises of course, from the fact that for such a γ^{n-1} of diameter $< \epsilon$, $\gamma^{n-1} \sim 0$ in D and hence in $E_n - B$. An important feature of this case is that the open set $E_n - \bar{D}$ is u.l.0-c., a property that does not necessarily hold under the hypothesis of Theorem 17. Even under the hypothesis of Theorem 17, however, $E_n - \bar{D}$ is u.l.i-c. for $1 \leqq i \leqq n - 2$, as follows easily from statements 2) and 3) of the conclusion and the fact, proved previously, that only a finite number of sets B_k^* are of diameter greater than a given positive number.

We can make strong use of the remarks of the preceding paragraph in extending Theorem 5 to the case of a set D as in Theorem 17. In the first place, for $1 \leqq i \leqq n - 2$, it may be shown that $p^i(D) \geqq p^{n-i-1}(E_n - \bar{D})$ by using the methods of the proof of Theorem 5. To show that $p^i(D) \leqq p^{n-i-1}(E_n - \bar{D})$, we proceed as follows: We add to D all those components of B that are points. We next modify our definition of B_k^* so as to accommodate the hypothesis that D is not necessarily connected; each B_k is on the boundary of one and only one component D_h of D; we let B_k^* consist of B_k and that component (there is only one, as shown above) of $E_n - \bar{D}_h$ that has its boundary in B_k. Obviously, as only a finite number of diameters of the sets B_k^* are $> \epsilon$ arbitrarily given, only a finite number of sets B_k^* can contain points of D. We further modify D, then, by adding to it every B_k^* which is such that its Betti numbers $p^i(B_k)$ are all zero, and such that B_k^* contains no points of D. The so modified D we call D'. That D' is open is easily seen. If it is identical with E_n we have no problem left, since this would mean than there are no linking i-cycles in $E_n - B$ and our duality is trivial. The boundary of D', then, consists of a finite number of components of B, say $B_1, B_2, \cdots, B_j$, each of which is a g.c. $(n - 1)$-m., and consequently the open set $D_1 = E_n - \bar{D}'$ is u.l.i-c. Therefore $p^i(D') \leqq p^{n-i-1}(E_n - \bar{D}')$, as is shown in the proof of Theorem 5.

Now $p^i(D) \leqq p^i(D')$. For a non-bounding cycle γ^i of D must link a B_k, $1 \leqq k \leqq j$, and as B_k is not in D', γ^i is non-bounding in D'. Similarly, i-cycles

linearly independent with respect to bounding in D are likewise linearly independent in D'. Hence $p^i(D) \leqq p^i(D')$.

Finally, $p^{n-i-1}(E_n - \bar{D}') \leqq p^{n-i-1}(E_n - \bar{D})$. For a non-bounding cycle γ^{n-i-1} of $E_n - \bar{D}'$ lies in the sets B_k^*, $1 \leqq k \leqq j$, and links the set $\sum_{k=1}^{j} B_k$; as the latter sets also lie in $E_n - D$, the γ^{n-i-1} is likewise non-bounding in $E_n - \bar{D}$. Similar remarks apply in the matter of the linear independence.

From the conjunction of the above relations, it follows that $p^i(D) \leqq p^{n-i-1}(E_n - \bar{D})$, and we have

THEOREM 18. *Let D be a bounded, uniformly locally i-connected* $(0 \leqq i \leqq n-2)$, *open subset of* E_n. *Then*

$$p^i(D) = p^{n-i-1}(E_n - \bar{D}), \qquad 1 \leqq i \leqq n-2.$$

In view of what we have shown in this section and in §2, we have also the following theorem:

THEOREM 19. *Let B be the boundary of a bounded, uniformly locally i-connected* $(0 \leqq i \leqq n-2)$, *open subset of* E_n. *Then the Betti numbers* $p^k(B)$ $(1 \leqq k \leqq n-2)$ *are all finite and the Poincaré duality holds:* $p^k(B) = p^{n-k-1}(B)$.

In conclusion, it may be of interest to point out one or two questions that have not been settled in the present paper. In the first place, it follows from Principal Theorems B_2 and C that a g.c.2-m. in E_3 is necessarily a manifold in the classical sense. It would be interesting to know if, for $n = 2$, a g.c.2-m. is always a manifold in the classical sense.

As to whether a g.c.*n*-m. is in general a manifold in the classical sense, the answer appears to be negative.[35] However, for the case of a g.c. $(n-1)$-m. in E_n, it seems quite likely that the two types of manifolds are identical.

It is my intention to consider, in a sequel to the present paper, 1) the finiteness of the Betti numbers and the validity of the Poincaré duality relation for the abstract g.c.*n*-m., and 2) the relations between the g.c.*n*-m. and the generalized manifolds, mentioned in the introduction, which have been introduced by Lefschetz and Čech.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON, NEW JERSEY

[35] In view of an example recently communicated to me by Dr. E. R. van Kampen.

Annals of Mathematics
Vol. 35, No. 4, October, 1934

THE GENERAL TOPOLOGICAL THEOREM OF DUALITY FOR CLOSED SETS[1]

By L. Pontrjagin

(Received August 8, 1934)

Introduction. In recent years there has appeared a series of papers devoted to topological theorems of duality. Leaving out the theorems of duality of Poincaré, which, by the way, are closely related to these questions, the general statement of the question is the following: Let M be a manifold and F one of its compact subsets. It is required to study the topological properties of the space $M - F$, starting from those of the set F. In particular, we may suppose that M is a Euclidean space or, what is essentially the same, a spherical manifold, F is homeomorphic to some complex, and the topological properties studied are the Betti numbers modulo two. In this form the problem was solved by Alexander.[2] Namely, he showed that the r-dimensional Betti number modulo two of the space $M - F$ equals the $(n - r - 1)$-dimensional Betti number of the complex F, where n is the dimensionality of the spherical manifold M.

This work of Alexander gave strong impetus to further development of the question, which went along the following lines:

1. The spherical manifold was replaced by a general manifold.
2. The assumption that F was homeomorphic to a complex was dropped.
3. Not only Betti groups modulo 2, but others as well, began to be studied.
4. The theory of linking was made use of in these studies; it is an application of the general theory of intersections developed by Lefschetz;[3] the last, together with the theories of homologies, is at present the basic well-developed apparatus of combinatorial topology. It is just because of their applications, it seems to me, that we owe the generalizations of the first three points. I will not dwell any more in detail on the history of this question; it is given in one of my papers.[4]

However, in all the work done so far, there remains unsolved the final question: letting F be an arbitrary compact sub set of M, where M is a Euclidean space, the reduction of the full Betti group of the space $M - F$ to the invariants of

---

[1] The results of this paper were reported in condensed form at the International Mathematical Congress in Zurich, 1932.

[2] Alexander, A proof and generalization of the Jordan-Brouwer Theorem, Trans. Amer. Math. Soc., 23 (1922), p. 333–349.

[3] Lefschetz, Intersections and transformations of complexes and manifolds, Trans. Amer. Math. Soc., 28 (1926), p. 1–39.

[4] Pontrjagin, Über den algebraischen Inhalt topologischer Dualitätssätze., Math. Ann., 105 (1931), p. 165–205.

the space F is not accomplished. It is to the solution of this question that the present paper is devoted. In it I limit myself to the case where the manifold M is a Euclidean space, since, with existing methods, the generalization to the case of an arbitrary manifold does not present any great difficulty.

For the solution of the question I have to introduce a new invariant of a compact metrizable space F in the form of a commutative continuous group connected with a definite dimension r, which I call the r-dimensional Betti group of the space F. In view of the introduction of continuous groups I have devoted considerable attention to their direct study.[5] I must note that the initiative for using continuous groups in combinatorial topology belongs not to me but to Alexander and Cohen[6] who study continuous groups, having in view their application to combinatorial topology; moreover they concentrate their attention on the reduction of such a group to a system of numerical invariants, which seems to me to be impossible in essence, and which leads them to an error (see appendix 1, T. T. G.).

In conclusion, I must express my gratitude to P. Alexandroff who has many times pointed out to me the desirability of solving the question to whose solution the present work is devoted.

1. Here I give the fundamental topological concepts in the somewhat generalized form in which I will have to use them in what follows.

Let K be an arbitrary finite or infinite complex, the set of whose r-dimensional oriented simplexes we shall denote by T_i^r, $i = 1, 2, \cdots$, and let

$$T_i^r \to \sum e_{ij}^r T_j^{r-1}$$

be the system of incidence relations for K, where e_{ij}^r is ± 1 or 0.

Further, let $\mathfrak{G}$ be an arbitrary discrete commutative group, written additively, with not more than a countable number of elements, which we shall make the basis for building the invariants of the complex K. The finite linear form $\sum_i \mathfrak{x}_i T_i^r$ with coefficients in $\mathfrak{G}$, $\mathfrak{x}_i \subset \mathfrak{G}$, we shall call an r-dimensional complex of K modulo $\mathfrak{G}$. The totality of all r-dimensional complexes of K modulo $\mathfrak{G}$ forms a group $\mathfrak{L}_{\mathfrak{G}}^r$ if we take for our composition the addition of linear forms. By the boundary of $\sum_i \mathfrak{x}_i T_i^r$ we shall mean the linear form $\sum_{i,j} \mathfrak{x}_i e_{ij}^r T_j^{r-1}$, which is an element of the group $\mathfrak{L}_{\mathfrak{G}}^{r-1}$. In this way, to each element of the group $\mathfrak{L}_{\mathfrak{G}}^r$ there corresponds, as its boundary, some element of the group $\mathfrak{L}_{\mathfrak{G}}^{r-1}$,

[5] To understand this paper it is necessary to know the first chapter of my paper, The theory of topological commutative groups, Ann. of Math., (2) 35, 1934, p. 361–370. In the following I will denote this last paper by T. T. G. It must be noted here that in T. T. G. there is a mistake, which was pointed out to me by J. v. Neumann, namely, on page 381 at the end of definition 1b, it is written "the intersection of ψ_i and φ_i contains only zero" where it should be written "the intersection of all ψ_i contains only zero".

[6] Alexander, J. W., and Cohen, L. W., A classification of the homology groups of compact spaces, Ann. of Math., (2) 33, 1932, p. 538–566.

and because of this correspondence we have a homomorphism of the group $\mathfrak{L}^r_{\mathfrak{G}}$ on a subgroup $\mathfrak{H}^{r-1}_{\mathfrak{G}}$ of the group $\mathfrak{L}^{r-1}_{\mathfrak{G}}$ with the divisor $\mathfrak{Z}^r_{\mathfrak{G}}$. The elements of the group $\mathfrak{Z}^r_{\mathfrak{G}}$ are the cycles modulo $\mathfrak{G}$, and the elements of $\mathfrak{H}^r_{\mathfrak{G}}$ are cycles homologous to zero modulo $\mathfrak{G}$. It is easy to see that $\mathfrak{H}^r_{\mathfrak{G}} \subset \mathfrak{Z}^r_{\mathfrak{G}}$. The factor group $\mathfrak{Z}^r_{\mathfrak{G}}/\mathfrak{H}^r_{\mathfrak{G}} = \mathfrak{B}^r_{\mathfrak{G}}$ we shall call the r-dimensional Betti group of the complex K modulo $\mathfrak{G}$; it is the common invariant of the complex K and the group $\mathfrak{G}$.

In the event $\mathfrak{G}$ is the additive group of integers, then $\mathfrak{B}^r_{\mathfrak{G}}$ is the full Betti group of the complex K. If, further, $\mathfrak{G}$ is the group of residues of the group of whole numbers modulo m, then $\mathfrak{B}^r_{\mathfrak{G}}$ is the Betti group of the complex K modulo m.

Let us now, in the previous consideration, substitute instead of a discrete group $\mathfrak{G}$ a continuous compact commutative group X, with the second countability axiom. Then, as before, every linear form of the type $\sum_{i=1}^{\infty} \alpha_i T^r_i$ with coefficients in X, $\alpha_i \subset$ X, will be called an r-dimensional complex modulo X. The totality of all r-dimensional complexes modulo X forms a group Λ^r_X, where we can now define continuity in Λ^r_X: namely, if U is a certain neighborhood of the zero of X and k is some positive whole number, then we define a neighborhood of the zero of Λ^r_x as the totality of all elements of the form $\sum_1^{\infty} \alpha_i T^r_i$, where $\alpha_j \subset U$, with $j \leqq k$. In exactly the same way as above we define the boundary of a complex modulo X, and we obtain a homomorphism of the group Λ^r_X on the subgroup H^{r-1}_X with the divisor Z^r_X. In this way the entire theory of homologies is applicable also for the modulus X. Making use of the diagonal process it is easy to show that Λ^r_X is compact and that, since the homomorphism of Λ^r_X in Λ^{r-1}_X is continuous, Z^r_X and H^{r-1}_X, and also $B^r_X = Z^r_X/H^r_X$ are compact groups.

2. It is known that the theorem of duality of Alexander does not hold for the full Betti group; that is, if K is a complex, located in Euclidean space R^n, the full r-dimensional Betti group $\mathfrak{B}^r(R^n - K)$ is not isomorphic to the full Betti group $\mathfrak{B}^{n-r-1}(K)$; moreover, the group $\mathfrak{B}^r(R^n - K)$ is in general not defined by the group $\mathfrak{B}^{n-r-1}(K)$. This was precisely the basic obstacle to the proof of invariance of the full Betti group of the domain complementary to a certain set F relative to various positions of F in Euclidean space. It turns out, however, that if the Betti groups of the space $R^n - K$ are constructed on the basis (see §1) of a certain discrete group $\mathfrak{G}$, that is, if we consider groups $\mathfrak{B}^r_{\mathfrak{G}}$, and if the groups of the complex K are constructed on the basis of the group of characters X of the group $\mathfrak{G}$ (see definition 1, T. T. G.), that is, we consider groups B^{n-r-1}_X, then these are already in one-to-one correspondence with each other. Thus, the groups B^{n-r-1}_X and $\mathfrak{B}^r_{\mathfrak{G}}$ form an orthogonal couple, that is, each of them is a group of characters of the other (see definitions 1′ and 3, T. T. G.). In particular, if $\mathfrak{G}$ is taken as the additive group of integers, then X transforms into the group K (see remark 5, T. T. G.) and since $\mathfrak{B}^r_{\mathfrak{G}}$ in this case transforms into the full Betti group $\mathfrak{B}^r(R^n - K)$, the invariance of the latter will be

proved. Precisely this case is of greatest interest to us. If we are to take, further, for the group $\mathfrak{G}$ the group of residues of integers modulo m, then the group X turns out also to be isomorphic to the group of residues modulo m; in this case the groups B^r_X and $\mathfrak{B}^r_{\mathfrak{G}}$ are the usual Betti groups modulo m, and as they are both finite, according to remark 3, T. T. G., we obtain their isomorphisms, that is, the theorem of duality of Alexander modulo m.

For the proof of the facts we have stated it is necessary to generalize certain basic concepts of combinatorial topology to the algebraic complexes not constituting linear forms relative to simplexes with integer coefficients but linear forms with coefficients that are elements of more general discrete and continuous groups. §§3 and 4 will be devoted to this, while §§5 and 6 will be devoted to the proof of the theorems of duality themselves.

3. Definition 1. Let X and $\mathfrak{G}$ be two groups forming an orthogonal couple (see definition 3, T. T. G.). Consider two algebraic polyhedral complexes α and $\mathfrak{a}$ of dimensions r and $n - r$ located in Euclidean space R^n in a general position.[3] Here complex α is a linear form with coefficients in X, while complex $\mathfrak{a}$ is a linear form with coefficients in $\mathfrak{G}$:

$$\alpha = \sum_{i=1}^{p} \alpha_i \epsilon_i, \qquad \mathfrak{a} = \sum_{j=1}^{q} \mathfrak{a}_j \epsilon'_j,$$

where ϵ_i and ϵ'_j are oriented simplexes, $\alpha_i \subset X$, $\mathfrak{a}_j \subset \mathfrak{G}$. Let us define $X(\alpha, \mathfrak{a})$, the index of intersection of complexes α and $\mathfrak{a}$, by putting

$$X(\alpha, \mathfrak{a}) = \sum_{i=1}^{p} \sum_{j=1}^{q} \alpha_i \mathfrak{a}_j X(\epsilon_i, \epsilon'_j).$$

By virtue of the fact that X and $\mathfrak{G}$ are orthogonal, the product $\alpha_i \mathfrak{a}_j$ is an element of the group K, while the index of intersection $X(\epsilon_i, \epsilon'_j)$ of the two simplexes is understood in the usual sense[3] and is an integer,[9] so that $X(\alpha, \mathfrak{a})$ is an element of the group K.

It is easy to see that if β and $\mathfrak{b}$ are two complexes, the first with coefficients in X, the second with coefficients in $\mathfrak{G}$, the dimensions of which are r and $n - r + 1$, then the usual formula[3] holds:

$$X(\dot{\beta}, \mathfrak{b}) = \epsilon(r) X(\beta, \dot{\mathfrak{b}}),$$

where $\epsilon(r) = +1$ or -1 depending upon the value of r, and $\dot{\beta}$ and $\dot{\mathfrak{b}}$ are respectively the boundaries of β and $\mathfrak{b}$.

Definition 2. Let γ and $\mathfrak{g}$ be two non-intersecting cycles in space R^n, of dimensions r and $n - r - 1$, the first of these with coefficients in X, the second with coefficients in $\mathfrak{G}$ (X and $\mathfrak{G}$, as always, form an orthogonal couple). Let $\mathfrak{d}$ be an arbitrary complex with boundary $\mathfrak{g}$. Let us define the coefficient of linking $\mathfrak{B}(\gamma, \mathfrak{g})$ of cycles γ and $\mathfrak{g}$ by putting $\mathfrak{B}(\gamma, \mathfrak{g}) = X(\gamma, \mathfrak{d})$. In this way $\mathfrak{B}(\gamma, \mathfrak{g})$ is an

element of the group K. The invariance of $\mathfrak{B}(\gamma, \mathfrak{g})$ is proved in the usual way.

A series of properties of $X(\alpha, \mathfrak{a})$ and $\mathfrak{B}(\gamma, \mathfrak{g})$ may be proved in the usual way;[7] I shall make use of them in what follows without giving proofs.

4. Definition 3. Let F be a compact metric space. We shall call an r-dimensional simplex on F the totality of $r + 1$ points on F (the vertices of the simplex). A simplex on F we shall call an ϵ-simplex if the totality of its vertices forms a set of diameter smaller than ϵ. A simplex on F we shall call oriented if, as usual, a certain order of its vertices is given to within an even permutation while an odd permutation gives an orientation of opposite sign. A finite linear form relative to oriented r-dimensional simplexes on F with coefficients in X we shall call a complex of dimension r on F, modulo X. The sum of two r-dimensional complexes on F is defined as the sum of the corresponding linear forms. If the complex consists of one simplex with coefficient α, we will define its boundary as the sum of all faces, taken with suitable orientation and with coefficient α. The boundary of an arbitrary complex on F we shall define as the sum of the boundaries of the terms in the linear form. A complex, the boundary of which is equal to zero, will be called a cycle. We shall call a complex an ϵ-complex if all its simplexes are ϵ-simplexes. A cycle is called ϵ-homologous to zero if it is the boundary of a certain ϵ-complex, $\zeta \underset{\epsilon}{\sim} 0$. Two cycles are ϵ-homologous to each other if their difference is ϵ-homologous to zero.

Definition 4. Let F be a compact metric space. We shall call the sequence $\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots$ of r-dimensional cycles on F, modulo X, an r-dimensional true cycle on F modulo X, if (*a*) ζ_m is an ϵ_m-cycle, (*b*) ζ_m and ζ_{m+1} are ϵ_m-homologous to each other, where $\lim \epsilon_m = 0$. The sum of two r-dimensional true cycles $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ and $Z' = (\zeta'_1, \zeta'_2, \cdots, \zeta'_m, \cdots)$ shall be called the true cycle

$$Z + Z' = (\zeta_1 + \zeta'_1, \zeta_2 + \zeta'_2, \cdots \zeta_m + \zeta'_m, \cdots).$$

A true cycle Z shall be called homologous to zero on F, $Z \sim 0$, if there exist numbers δ_m, $\lim \delta_m = 0$, such that $\zeta_m \underset{\delta_m}{\sim} 0$. Two true cycles are homologous to each other if their difference is homologous to zero.

Definition 5. Let F be a compact subset of n-dimensional Euclidean space. Further, let $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ be an r-dimensional true cycle on F modulo X and $\mathfrak{z}$ an $n - r - 1$-dimensional cycle in space $R^n - F$ modulo $\mathfrak{G}$ (X and $\mathfrak{G}$ are orthogonal groups). Let us define the coefficient of linking $\mathfrak{B}(Z, \mathfrak{z})$ of cycles Z and $\mathfrak{z}$. Each simplex of ζ_m is given only by its vertices. Let us realize it in the form of the polyhedral geometrical simplex, taking it with the same coefficient with which it enters in ζ_m and let us form the sum of all the simplexes entering in ζ_m, in this way obtaining a cycle $\bar{\zeta}_m$ giving a geometrical realization of ζ_m. Let us define $\mathfrak{B}(Z, \mathfrak{z})$ as $\mathfrak{B}(\bar{\zeta}_m, \mathfrak{z})$ for sufficiently large m;

[7] For the fundamental concepts of combinatorial topology, see the book of Lefschetz, Topology.

obviously, for sufficiently large m, $\mathfrak{B}(\bar{\zeta}_m, \mathfrak{z})$ is independent of m, since all the cycles of $\bar{\zeta}_m$ with sufficiently large m are homologous to each other outside the cycle $\mathfrak{z}$. Thus, $\mathfrak{B}(Z, \mathfrak{z})$ is an element of the group K.

DEFINITION 6. Let us define an r-dimensional Betti group $\mathrm{B}_{\mathrm{X}}^r$ modulo X of a compact metric space F. Each class α of true r-dimensional cycles homologous to each other in pairs modulo X on F is an element of the group $\mathrm{B}_{\mathrm{X}}^r$. If Z is a certain true cycle of class α and Z$'$ is a certain true cycle of class α', we shall call the sum $\alpha + \alpha'$, a class β which contains the true cycle Z + Z$'$. It is easy to see that the sum $\alpha + \alpha'$ defined in this way does not depend upon the arbitrariness of choice of Z and Z$'$, but depends only upon the elements α and α' themselves. In this way the algebraic operations in $\mathrm{B}_{\mathrm{X}}^r$ are defined. Let us now define the topological operations in $\mathrm{B}_{\mathrm{X}}^r$ in order to make this group continuous; for this it is sufficient to give a full system of neighborhoods of zero in $\mathrm{B}_{\mathrm{X}}^r$. Let us define a neighborhood V of zero in $\mathrm{B}_{\mathrm{X}}^r$ as a function of a certain positive whole number k and a certain neighborhood U of zero of the group X. The element α of the group $\mathrm{B}_{\mathrm{X}}^r$ we shall regard as belonging to V if: (a) there exists a true cycle Z of class α, Z $= (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$, where ζ_m is an ϵ_m-cycle, and $\zeta_m \underset{\epsilon_m}{\sim} \zeta_{m+1}$, such that $\epsilon_m < 1/k$, for all m; (b) $\alpha_i, i = 1, 2, \cdots l$, being a totality of all elements of the group X which are coefficients of simplexes of ζ_1, any element of X of the form $\sum_{i=1}^{l} a_i\alpha_i$ belongs to U if a_i are arbitrary integers whose absolute values do not exceed unity. As k runs over all positive integral values and U a full system of neighborhoods of zero of the group X, V ranges over a full system of neighborhoods of zero of the group $\mathrm{B}_{\mathrm{X}}^r$. Since there exists a countable full system of neighborhoods of zero in X and the set of integers is countable, the group $\mathrm{B}_{\mathrm{X}}^r$ also admits a countable full system of neighborhoods of zero. Let us now define the neighborhood W of an arbitrary element $\alpha \subset \mathrm{B}_{\mathrm{X}}^r$ as a function of a certain neighborhood of zero V. We shall consider that $\beta \subset W$ if $\beta - \alpha \subset V$. When V ranges over a full system of neighborhoods of zero, W runs over a full system of neighborhoods of α. Since the full system of neighborhoods of zero in $\mathrm{B}_{\mathrm{X}}^r$, by what was proved, is countable, the full system of neighborhoods of α is also countable; consequently, the first axiom of countability in $\mathrm{B}_{\mathrm{X}}^r$ is fulfilled.

THEOREM 1. *A Betti group* $\mathrm{B}_{\mathrm{X}}^r = \mathrm{B}$ *modulo* X *of a compact metric space* F *is compact and satisfies the second axiom of countability.*

PROOF: Let $A_1, A_2, \cdots, A_m, \cdots$ be a sequence of complexes giving a projective spectrum F,[8] and let Π_m be the simplicial reflection of A_{m+1} in A_m arising from the projective spectrum. It is easy to see that there exists a sequence of positive numbers $\epsilon_1, \epsilon_2, \cdots, \epsilon_m, \cdots$ such that: a) $\lim \epsilon_m = 0$; b) the simplexes A_m have a diameter less than ϵ_m; c) if ζ is a cycle of A_{m+1}, then ζ and $\Pi_m(\zeta)$ are ϵ_m-homologous on F.

---

[8] Alexandroff, Gestalt und Lage abgeschlossener Mengen, Ann. of Math. (2) 30, 1928, p. 107. But instead of a projective spectrum, it is possible here to use any approximating sequence of complexes.

element of the group K. The invariance of $\mathfrak{B}(\gamma, \mathfrak{g})$ is proved in the usual way.

A series of properties of $X(\alpha, \mathfrak{a})$ and $\mathfrak{B}(\gamma, \mathfrak{g})$ may be proved in the usual way;[7] I shall make use of them in what follows without giving proofs.

4. Definition 3. Let F be a compact metric space. We shall call an r-dimensional simplex on F the totality of $r + 1$ points on F (the vertices of the simplex). A simplex on F we shall call an ϵ-simplex if the totality of its vertices forms a set of diameter smaller than ϵ. A simplex on F we shall call oriented if, as usual, a certain order of its vertices is given to within an even permutation while an odd permutation gives an orientation of opposite sign. A finite linear form relative to oriented r-dimensional simplexes on F with coefficients in X we shall call a complex of dimension r on F, modulo X. The sum of two r-dimensional complexes on F is defined as the sum of the corresponding linear forms. If the complex consists of one simplex with coefficient α, we will define its boundary as the sum of all faces, taken with suitable orientation and with coefficient α. The boundary of an arbitrary complex on F we shall define as the sum of the boundaries of the terms in the linear form. A complex, the boundary of which is equal to zero, will be called a cycle. We shall call a complex an ϵ-complex if all its simplexes are ϵ-simplexes. A cycle is called ϵ-homologous to zero if it is the boundary of a certain ϵ-complex, $\zeta \underset{\epsilon}{\sim} 0$. Two cycles are ϵ-homologous to each other if their difference is ϵ-homologous to zero.

Definition 4. Let F be a compact metric space. We shall call the sequence $\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots$ of r-dimensional cycles on F, modulo X, an r-dimensional true cycle on F modulo X, if (a) ζ_m is an ϵ_m-cycle, (b) ζ_m and ζ_{m+1} are ϵ_m-homologous to each other, where $\lim \epsilon_m = 0$. The sum of two r-dimensional true cycles $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ and $Z' = (\zeta'_1, \zeta'_2, \cdots, \zeta'_m, \cdots)$ shall be called the true cycle

$$Z + Z' = (\zeta_1 + \zeta'_1, \zeta_2 + \zeta'_2, \cdots \zeta_m + \zeta'_m, \cdots).$$

A true cycle Z shall be called homologous to zero on F, $Z \sim 0$, if there exist numbers δ_m, $\lim \delta_m = 0$, such that $\zeta_m \underset{\delta_m}{\sim} 0$. Two true cycles are homologous to each other if their difference is homologous to zero.

Definition 5. Let F be a compact subset of n-dimensional Euclidean space. Further, let $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ be an r-dimensional true cycle on F modulo X and $\mathfrak{z}$ an $n - r - 1$-dimensional cycle in space $R^n - F$ modulo $\mathfrak{G}$ (X and $\mathfrak{G}$ are orthogonal groups). Let us define the coefficient of linking $\mathfrak{B}(Z, \mathfrak{z})$ of cycles Z and $\mathfrak{z}$. Each simplex of ζ_m is given only by its vertices. Let us realize it in the form of the polyhedral geometrical simplex, taking it with the same coefficient with which it enters in ζ_m and let us form the sum of all the simplexes entering in ζ_m, in this way obtaining a cycle $\bar{\zeta}_m$ giving a geometrical realization of ζ_m. Let us define $\mathfrak{B}(Z, \mathfrak{z})$ as $\mathfrak{B}(\bar{\zeta}_m, \mathfrak{z})$ for sufficiently large m;

[7] For the fundamental concepts of combinatorial topology, see the book of Lefschetz, Topology.

obviously, for sufficiently large m, $\mathfrak{V}(\mathfrak{z}_m, \mathfrak{z})$ is independent of m, since all the cycles of $\mathfrak{z}_m$ with sufficiently large m are homologous to each other outside the cycle $\mathfrak{z}$. Thus, $\mathfrak{V}(Z, \mathfrak{z})$ is an element of the group K.

DEFINITION 6. Let us define an r-dimensional Betti group B^r_X modulo X of a compact metric space F. Each class α of true r-dimensional cycles homologous to each other in pairs modulo X on F is an element of the group B^r_X. If Z is a certain true cycle of class α and Z$'$ is a certain true cycle of class α', we shall call the sum $\alpha + \alpha'$, a class β which contains the true cycle Z + Z$'$. It is easy to see that the sum $\alpha + \alpha'$ defined in this way does not depend upon the arbitrariness of choice of Z and Z$'$, but depends only upon the elements α and α' themselves. In this way the algebraic operations in B^r_X are defined. Let us now define the topological operations in B^r_X in order to make this group continuous; for this it is sufficient to give a full system of neighborhoods of zero in B^r_X. Let us define a neighborhood V of zero in B^r_X as a function of a certain positive whole number k and a certain neighborhood U of zero of the group X. The element α of the group B^r_X we shall regard as belonging to V if: (a) there exists a true cycle Z of class α, $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$, where ζ_m is an ϵ_m-cycle, and $\zeta_m \underset{\epsilon_m}{\sim} \zeta_{m+1}$, such that $\epsilon_m < 1/k$, for all m; (b) $\alpha_i, i = 1, 2, \cdots l$, being a totality of all elements of the group X which are coefficients of simplexes of ζ_1, any element of X of the form $\sum_{i=1}^{l} a_i\alpha_i$ belongs to U if a_i are arbitrary integers whose absolute values do not exceed unity. As k runs over all positive integral values and U a full system of neighborhoods of zero of the group X, V ranges over a full system of neighborhoods of zero of the group B^r_X. Since there exists a countable full system of neighborhoods of zero in X and the set of integers is countable, the group B^r_X also admits a countable full system of neighborhoods of zero. Let us now define the neighborhood W of an arbitrary element $\alpha \subset B^r_X$ as a function of a certain neighborhood of zero V. We shall consider that $\beta \subset W$ if $\beta - \alpha \subset V$. When V ranges over a full system of neighborhoods of zero, W runs over a full system of neighborhoods of α. Since the full system of neighborhoods of zero in B^r_X, by what was proved, is countable, the full system of neighborhoods of α is also countable; consequently, the first axiom of countability in B^r_X is fulfilled.

THEOREM 1. *A Betti group* $B^r_X = B$ *modulo* X *of a compact metric space* F *is compact and satisfies the second axiom of countability.*

PROOF: Let $A_1, A_2, \cdots, A_m, \cdots$ be a sequence of complexes giving a projective spectrum F,[8] and let Π_m be the simplicial reflection of A_{m+1} in A_m arising from the projective spectrum. It is easy to see that there exists a sequence of positive numbers $\epsilon_1, \epsilon_2, \cdots, \epsilon_m, \cdots$ such that: a) $\lim \epsilon_m = 0$; b) the simplexes A_m have a diameter less than ϵ_m; c) if ζ is a cycle of A_{m+1}, then ζ and $\Pi_m(\zeta)$ are ϵ_m-homologous on F.

[8] Alexandroff, Gestalt und Lage abgeschlossener Mengen, Ann. of Math. (2) 30, 1928, p. 107. But instead of a projective spectrum, it is possible here to use any approximating sequence of complexes.

It is likewise not difficult to show that from every class of homologies α, a true cycle $Z = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ can be chosen in such a way that Z_m is a cycle of A_m, and ζ_{m+1} and ζ_m are ϵ_m-homologous to each other on F. Let $\beta_1, \beta_2, \cdots \beta_i, \cdots$ be an arbitrary sequence of elements of B. We shall show that this sequence has a limit element. To do this we choose from β_i a true cycle $Z_i = (\zeta_{1,i}, \zeta_{2,i}, \cdots, \zeta_{m,i} \cdots)$ of the type indicated. Since the group of cycles modulo X of the complex A_m is compact for every m, we can make use of the diagonal process to choose from the sequence $\bar{Z}_i$ the sequence Z_{i_j} such that for every m, $\lim_{j\to\infty} \zeta_{m,i_j} = \bar{\zeta}_m$, where $\bar{\zeta}_m$ is a cycle of A_m and moreover $\bar{\zeta}_m$ and $\bar{\zeta}_{m+1}$ are ϵ_m-homologous to each other on F. In this way the sequence of cycles $\bar{\zeta}_1, \bar{\zeta}_2, \bar{\zeta}_3, \cdots, \bar{\zeta}_m, \cdots$ forms a true cycle $\bar{Z}$. We shall show that $\lim_{j\to\infty} \beta_{i_j} = \bar{\beta}$, where $\bar{\beta}$ is the class of homologies containing $\bar{Z}$. Let V be an arbitrary neighborhood of zero in B, and let U and k be that neighborhood of zero and that integer, which by definition 6 determine the neighborhood V. Let m' be sufficiently large that for $n \geqq m'$, $\epsilon_n < 1/k$. Let us designate by s the number of r-dimensional simplexes of $A_{m'}$. And let us further choose a neighborhood G of zero in X such that if $\alpha_i \subset G$ then $\sum_{i=1}^{s} a_i\alpha_i \subset U$, where a_i are integers and $|a_i| \leqq 1$. Let Z^r be the group of r-dimensional cycles of the complex A_m, and let W be a neighborhood of zero in Z^r such that W contains only cycles the coefficients of all of whose simplexes are contained in G. Since $\lim_{j\to\infty} z_{m',i_j} = \bar{\zeta}_{m'}$, there exists a p sufficiently large that with $j > p$, $\bar{\zeta}_{m'} - \zeta_{m',i_j} \subset W$. If we denote by γ_i, $i = 1, 2, \cdots, l$, the totality of coefficients in the simplexes of the cycle $\bar{\zeta}_{m'} - z_{m',i_j}$ where $l \leqq s$, then from the construction of W we shall have $\gamma_i \subset G$, and from the definition of G, $\sum_{i=1}^{l} a_i\gamma_i \subset U$ for arbitrary integers a_i, whose absolute values do not exceed unity.

Since the obviously true cycle $(\bar{\zeta}_{m'} - \zeta_{m',i_j}), (\bar{\zeta}_{m'+1} - \zeta_{m'+1,i_j}), \cdots$, belongs to the class $\bar{\beta} - \beta_{i_j}$, it follows from definition 6 that for $j > p$, $\bar{\beta} - \beta_{i_j} \subset V$. In this way we have $\lim_{j\to\infty} \beta_{i_j} = \bar{\beta}$ but the sequence β_{i_j} is a subsequence of an arbitrary sequence β_i. Hence the compactness of B is established.

Now we shall show that if B is an arbitrary compact group for which the first axiom of countability holds, then the second axiom of countability holds likewise. We shall show first of all that B contains a denumerable everywhere-dense set N. It is easily seen that in every neighborhood V of zero in B there exists a finite set M such that for every $\alpha \subset$ B there is a $\mu \subset M$ such that $\alpha - \mu \subset V$. Taking a complete denumerable system of neighborhoods of zero in B and constructing the corresponding sets M, and then summing them, we shall get for the sum a set N which is everywhere dense on B. The totality of the complete denumerable systems of neighborhoods for all the elements of N now forms a complete denumerable system of neighborhoods for the whole B. In fact, let G be any domain of B and let $\alpha \subset G$. Then there exists a neighborhood of zero V such that if $\beta \subset V$, $\gamma \subset V$, then $\alpha + \beta - \gamma \subset G$. Furthermore we can find a $\delta \subset V$ such that $\alpha - \delta = \nu \subset N$. The totality of all the

elements of the form $\nu + \eta$, where $\eta \subset V$, constitute the neighborhood W of the point ν with $\alpha \subset W \subset G$. In this way for every domain G and element $\alpha \subset G$ there exists a neighborhood W of our denumerable system such that $\alpha \subset W \subset G$, that is, the system is complete.

5. Theorem of duality for a complex.

THEOREM 2. *Let the complex K be polyhedrally situated in R^n, let $\mathrm{B}^r_{\mathrm{X}}$ be an r-dimensional Betti group of K modulo* X *and let $\mathfrak{B}^s_{\mathfrak{G}}$ be an s-dimensional Betti group in the space $(R^n - K)$ modulo $\mathfrak{G}$.* (X *and* $\mathfrak{G}$ *are orthogonal.) Then $\mathrm{B}^r_{\mathrm{X}}$ and $\mathfrak{B}^{n-r-1}_{\mathfrak{G}}$ are orthogonal, with the product of $\alpha \subset B^r_{\mathrm{X}}$ and $\mathfrak{a} \subset \mathfrak{B}^{n-r-1}_{\mathfrak{G}}$ determined as the coefficient of linking (see definition 2) of some cycle of class α with a cycle of class $\mathfrak{a}$.*[9]

PROOF: Let us suppose that K is composed of simplexes of some subdivision of R^n. Let T^r_i, $i = 1, 2, \cdots a_r$, be the totality of r-dimensional oriented simplexes of K. Let us denote further by S^s_j, $j = 1, 2, \cdots, a_{n-s}$, the totality of s-dimensional barycentric stars of R^n which intersect K. It is obvious that the number of r-dimensional simplexes of K is equal to the number of $n - r$-dimensional stars of S^{n-r}_j, which is foreseen in the nomenclature. Moreover, the notation can be chosen in such a way that the index of intersection $\mathrm{X}(T^r_i, S^{n-r}_j) = \delta^j_i$, where $\delta^j_i = 0$ if $i \neq j$ and $\delta^i_i = 1$.

Let us denote by Λ^r the group of all linear forms relative to all r-dimensional simplexes of the complex K with coefficients in X, that is, the group of r-dimensional subcomplexes K modulo X. To every element in Λ^r corresponds, as its boundary, some element of Λ^{r-1} (see §1). In this way we have a homomorphic reflection of the group Λ^r on the subgroup H^{r-1} of the group Λ^{r-1} with the divisor Z^r. The elements of Z^r are cycles, and the elements of H^{r-1} are cycles homologous to zero. Analogously we designate by $\mathfrak{L}^s$ the group of all the linear forms of the type $\sum_{j=i}^{a_{n-s}} \mathfrak{x}_j S^s_j$, where $\mathfrak{x}_j \subset \mathfrak{G}$; in this way $\mathfrak{L}^s$ is a group of complexes modulo $\mathfrak{G}$. To every element $\mathfrak{a}$ of $\mathfrak{L}^s$ corresponds, as its boundary, some linear form relative to barycentric stars; we shall choose only those terms of the form which are composed of S^{s-1}_j and we shall call their sum a relative boundary of $\mathfrak{a}$. In this way to every element of $\mathfrak{L}^s$ corresponds, as its relative boundary, some element of $\mathfrak{L}^{s-1}$, and we get a homomorphic reflection of the group $\mathfrak{L}^s$ on the subgroup $\mathfrak{H}^{s-1}$ of the group $\mathfrak{L}^{s-1}$, with the divisor $\mathfrak{Z}^s$. It is easy to see that $\mathfrak{H}^s \subset \mathfrak{Z}^s$.

We shall now establish the law of multiplication between the elements of Λ^r and $\mathfrak{L}^{n-r}$. We shall suppose that if $\alpha \subset \Lambda^r$ and $\mathfrak{a} \subset \mathfrak{L}^{n-r}$ then $\alpha\,\mathfrak{a} = \mathrm{X}(\alpha, \mathfrak{a})$ (see definition 1). We shall show that the groups Λ^r and $\mathfrak{L}^{n-r}$, because of this law of multiplication, form an orthogonal pair. That the multiplication is distributive and continuous is obvious. Let us show that $(\Lambda^r, \mathfrak{L}^{n-r}) = (0)$ (see

[9] The proof of this theorem is taken from my paper Über den algebraischen Inhalt . . . . (see note 3), p. 186, II Formulierung und . . . .

definition 3, T. T. G.). Let $0 \neq \alpha \subset \Lambda^r$, $\alpha = \sum_{i=1}^{a_r} \alpha_i T_i^r$, with at least one of the coefficients α_i different from zero. Let it be α_ϵ. Since X and $\mathfrak{G}$ are orthogonal there exists an $\mathfrak{x} \subset \mathfrak{G}$ such that $\alpha_\epsilon \mathfrak{x} \neq 0$.

Let $\mathfrak{a} = \mathfrak{x} S_\epsilon^{n-r}$; we shall have $\alpha\, \mathfrak{a} = \alpha_\epsilon \mathfrak{x} \neq 0$. In exactly the same way we can show that $(\mathfrak{L}^{n-r}, \Lambda^r) = (0)$.

We also show that $(\Lambda^r, \mathfrak{H}^{n-r}) = \mathrm{Z}^r$ and $(\mathfrak{L}^{n-r}, \mathrm{H}^r) = \mathfrak{Z}^{n-r}$. Let $\alpha \subset \mathrm{Z}^r$ and $\mathfrak{a} \subset \mathfrak{H}^{n-r}$. There exists an element $\mathfrak{c} \subset \mathfrak{L}^{n-r+1}$ whose relative boundary is $\mathfrak{a}$, that is, $\mathfrak{c} \rightarrow \mathfrak{a} + \mathfrak{b}$, where $\mathfrak{b}$ does not intersect K . $\alpha\, \mathfrak{a} = \mathrm{X}(\alpha, \mathfrak{a}) = \mathrm{X}(\alpha, \mathfrak{a} + \mathfrak{b}) = 0$, since the index of intersection of two cycles in R^n is equal to zero. In this way $(\Lambda^r, \mathfrak{H}^{n-r}) \supset \mathrm{Z}^r$. Now let $\alpha \subset \Lambda^r$, $\alpha \not\subset \mathrm{Z}^r$, then $\alpha \rightarrow \beta \neq 0$, $\beta \subset \Lambda^{r-1}$. Since Λ^{r-1} and $\mathfrak{L}^{n-r+1}$ are orthogonal, there exists a $\mathfrak{c} \subset \mathfrak{L}^{n-r+1}$ such that $\beta\, \mathfrak{c} \neq 0$. Let $\mathfrak{c} \rightarrow \mathfrak{a} + \mathfrak{b}$, $\mathfrak{a} \subset \mathfrak{H}^{n-r}$, $\mathfrak{b}$ does not intersect K. Then $\alpha\, \mathfrak{a} = \mathrm{X}(\alpha, \mathfrak{a} + \mathfrak{b}) = \pm\, \mathrm{X}(\beta, \mathfrak{c}) = \pm\, \beta\, \mathfrak{c} \neq 0$. In this way $(\Lambda^r, \mathfrak{H}^{n-r}) = \mathrm{Z}^r$. Quite analogously we can prove that $(\mathfrak{L}^{n-r}, \mathrm{H}^r) = \mathfrak{Z}^{n-r}$.

From what we have just shown together with theorems 2 and 4 (T. T. G.) it follows that

$$(\mathfrak{L}^{n-r}, \mathrm{Z}^r) = \mathfrak{H}^{n-r}, \qquad (\Lambda^r, \mathfrak{Z}^{n-r}) = \mathrm{H}^r.$$

From which, taking into consideration that Z^r and $\mathfrak{Z}^{n-r}$ form a pair (see definition 3, T. T. G.), we conclude that

$$(\mathrm{Z}^r, \mathfrak{Z}^{n-r}) = \mathrm{H}^r, \qquad (\mathfrak{Z}^{n-r}, \mathrm{Z}^r) = \mathfrak{H}^{n-r}.$$

But from this follows the orthogonality of the groups $\mathrm{B}_{\mathrm{X}}^r = \mathrm{Z}^r/\mathrm{H}^r$ and $\mathfrak{Z}^{n-r}/\mathfrak{H}^{n-r}$. The first of these is a Betti group of the complex K, the second is isomorphic with the Betti group $\mathfrak{B}^{n-r-1}$ in the space $R^n - K$. This last statement is demonstrated as follows: let $\mathfrak{z} \subset \mathfrak{Z}^{n-r}$, $\mathfrak{z} \rightarrow \mathfrak{z}'$; the boundary $\mathfrak{z}'$ of the complex $\mathfrak{z}$ is a cycle in the space $R^n - K$ and determines some element of the group $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$; at the same time if $\mathfrak{y} \subset \mathfrak{Z}^{n-r}$, and $\mathfrak{z} - \mathfrak{y} \subset \mathfrak{H}^{n-r}$, then the boundaries of $\mathfrak{z}$ and $\mathfrak{y}$ will be cycles homologous to each other in R^n. Conversely if $\mathfrak{z}$ and $\mathfrak{y}$ belong to $\mathfrak{Z}^{n-r}$ and if their boundaries are homologous to each other in R^n, then $\mathfrak{z} - \mathfrak{y} \subset \mathfrak{H}^{n-r}$. Moreover, for every $(n - r - 1)$-dimensional cycle of $R^n - K$, a cycle can be found, homologous to it, which forms the boundary of some element of $\mathfrak{Z}^{n-r}$. In this way the isomorphism between the groups $\mathfrak{Z}^{n-r}/\mathfrak{H}^{n-r}$ and $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ is established. It is obvious that by going from elements of the group $\mathfrak{Z}^{n-r}$ to their boundaries, the indices of their intersections with the elements of Z^r go over into the coefficients of linking. Hence the theorem is proved.

6. The general theorem of duality for a compact set F located in R^n.

FUNDAMENTAL THEOREM. *Let F be in R^n, let $\mathrm{B}_{\mathrm{X}}^r$ be an r-dimensional Betti group of F modulo* X, *and let $\mathfrak{B}_{\mathfrak{G}}^s$ be an s-dimensional Betti group modulo $\mathfrak{G}$ of the space $R^n - F$.* (X *and $\mathfrak{G}$ are orthogonal.*) *Then the groups $\mathrm{B}_{\mathrm{X}}^r$ and $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ are orthogonal, the product of $\alpha \subset \mathrm{B}_{\mathrm{X}}^r$ and $\mathfrak{a} \subset \mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ being determined as the coefficient of linking of some true cycle of class α with a cycle of class $\mathfrak{a}$.*

COROLLARY: *It follows from this Theorem, in view of Theorem 5, T. T. G., that each of the groups* $\mathrm{B}_{\mathrm{X}}^{r}$ *and* $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ *is the group of characters of the other, that is, each of them is determined uniquely by the other.*

PROOF OF FUNDAMENTAL THEOREM. The law of multiplication formulated in the theorem is obviously distributive. Let us show now that it satisfies the condition of continuity of multiplication. To do this, it is sufficient to show that if $\mathfrak{a} \subset \mathfrak{B}_{\mathfrak{G}}^{n-r-1}$, then for any neighborhood J of zero of the group K, there is always a neighborhood V of zero in the group $\mathrm{B}_{\mathrm{X}}^{r}$, such that if $\alpha \subset V$, then $\alpha\,\mathfrak{a} \subset J$. Let $\mathfrak{z}$ be a cycle of the class $\mathfrak{a}$ and let $\mathfrak{b} \to \mathfrak{z}$. Let $\mathfrak{x}_i$, $i = 1, 2, \cdots, p$, be the totality of all the coefficients in the simplexes of the complex $\mathfrak{b}$. Let us denote by U a sufficiently small neighborhood of zero in X, such that if $\beta_i \subset U$, then $\sum_{j=1}^{p} \beta_j\, \mathfrak{x}_j \subset J$. Further let K be such a small polyhedral neighborhood of F that K does not intersect $\mathfrak{z}$; let ϵ be a sufficiently small number so that every ϵ-complex on F lies in K, and let k be a whole number so large that $1/k < \epsilon$. Let V be the neighborhood of zero in $\mathrm{B}_{\mathrm{X}}^{r}$ which is determined by the neighborhood U and the number k according to definition 6. We shall show that if $\alpha \subset V$, then $\alpha\,\mathfrak{a} \subset J$. Let $\mathrm{Z} = (\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots)$ be a true cycle of class α such that, by definition 6, α belongs to the neighborhood V. Since by force of this condition all the cycles ζ_m, and all the complexes which establish homologies between them, belong to K, we have from definition 5, $\mathfrak{V}(\mathrm{Z} \cdot \mathfrak{z}) = \mathfrak{V}(\zeta_1, \mathfrak{z})$. Let α_i, $i = 1, 2, \cdots l$, be the totality of all the coefficients of the simplexes of the cycle ζ_1; then we have from definitions 1 and 2, $\mathfrak{V}(\mathrm{Z}, \mathfrak{z}) = \mathrm{X}(\zeta_1, \mathfrak{b}) = \sum_{i=1}^{l} \sum_{j=1}^{p} \alpha_i\, \mathfrak{x}_j\, a_{ij}$, where a_{ij} are integers whose absolute values do not exceed unity. By definition 6, $\sum_{i=1}^{l} a_{ij}\, \alpha_i = \beta_j \subset U$. In this way from the construction of U, $\sum_{i=1}^{l} \sum_{j=1}^{p} \alpha_i\, \mathfrak{x}_j\, a_{ij} = \sum_{j=1}^{p} \beta_j\, \mathfrak{x}_j \subset J$. From the multiplication law $\mathrm{B}_{\mathrm{X}}^{r}$ and $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ form a pair (see definition 3, T. T. G.).

We shall now prove that $\mathrm{B}_{\mathrm{X}}^{r}$ and $\mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ are orthogonal. Let $0 \neq \alpha \subset \mathrm{B}_{\mathrm{X}}^{r}$, we shall show that there is an element $\mathfrak{a} \subset \mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ such that $\alpha\,\mathfrak{a} \neq 0$. Let $(\zeta_1, \zeta_2, \cdots, \zeta_m, \cdots) = \mathrm{Z}$ be a true cycle belonging to the class α. Since $\alpha \neq 0$, there exists a polyhedral neighborhood K of the set F such that the cycles ζ_m are not homologous to zero in it. We can suppose without loss of generality that all the cycles ζ_m and all the complexes, which establish homologies between them, are in K. Since $\zeta_1 \subset K$, and is not homologous to zero in K, it follows from Theorem 2 that there is a cycle $\mathfrak{z} \subset R^n - K$ such that $V(\zeta_1, \mathfrak{z}) \neq 0$. Let $\mathfrak{a} \subset \mathfrak{B}_{\mathfrak{G}}^{n-r-1}$ be a class to which $\mathfrak{z}$ belongs; then $\alpha\,\mathfrak{a} \neq 0$. Hence $(\mathrm{B}_{\mathrm{X}}^{r}, \mathfrak{B}_{\mathfrak{G}}^{n-r-1}) = (0)$. We shall now show that if $0 \neq \mathfrak{a} \subset \mathfrak{B}_{\mathfrak{G}}^{n-r-1}$, there exists an $\alpha \subset \mathrm{B}_{\mathrm{X}}^{r}$ such that $\alpha\,\mathfrak{a} \neq 0$. Let $\mathfrak{z}$ be a cycle of class $\mathfrak{a}$, and let $K_1, K_2, \cdots, K_m, \cdots$ be a decreasing sequence of complexes containing F whose intersection coincides with F. We also suppose that $\mathfrak{z} \subset R^n - K_1$. Let g be the order of a, and let γ be some element of K of order g. From Theorem 2 and Lemma 8, T. T. G., there exists in K_m a cycle ζ_m such that $V(\zeta_m, \mathfrak{z}) = \gamma$. Let ζ_m^j for $j \geqq m$ be a cycle composed of simplexes K_m homologous in K_m with the cycle ζ_j. Since the group of cycles of the complex K_m is compact, we

can choose from the sequence ζ_i^j $(j = 1, 2, \cdots)$ a converging subsequence $\zeta_1^{j_k}$ $(k = 1, 2, \cdots)$. In the same way from the sequence $\zeta_2^{j_k}$ a convergent subsequence can be chosen. Continuing this process further and using the diagonal process we shall arrive at a sequence of numbers $i_1, i_2, \cdots, i_q, \cdots$ such that the sequence $\zeta_m^{i_q}$, with m fixed, converges for every m to some cycle $\bar{\zeta}_m$. It is obvious that $\bar{\zeta}_{m+1} \sim \bar{\zeta}_m$ in K_m, and that $V(\bar{\zeta}_m, \mathfrak{z}) = \gamma$. Pushing the vertices of the cycles $\bar{\zeta}_m$ on to F, we obtain a true cycle Z belonging to some class $\alpha \subset \mathrm{B}_X^r$ with $\alpha\, \mathfrak{a} = \gamma$. Hence $(\mathfrak{B}_{\mathfrak{S}}^{n-r-1}, \mathrm{B}_X^r) = (0)$ and the theorem is completely proved.

Moscow, U. S. S. R.

ERRATA

Page 119. line 23: replace each "LC" by "LC^∞".
Page 126. line 5 from the bottom, replace "convex sets of H" by "spheres."
line 4 from the bottom, suppress "convex."
line 3 from the bottom suppress the whole line.
Page 127. line 13 from the bottom, replace K^* by R^*.

Owing to unavoidable circumstances (lost in transit abroad) the paper "On local properties of closed sets" by P. Alexandroff due to appear in the last number of volume 35, had to be postponed to the January number of the next volume.

9

Egleston

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

S. LEFSCHETZ J. VON NEUMANN

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE FOR ADVANCED STUDY

AND

SECOND SERIES, VOL. 35, NO. 1

JANUARY, 1934

PRINCETON UNIVERSITY PRESS

PRINCETON, N. J.

1934

ANNALS OF MATHEMATICS

Published at Princeton, N. J. Subscription price, $5, a volume (four numbers) in advance. Single copies $1.50. An arrangement has been made with the Mathematical Association of America in accordance with which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulae, and must be in final form with all references filled in.

Authors receive gratis 100 reprints of each article, postage prepaid.

Application has been made for entry as second class matter at the post office at Baltimore, Md.

Made in United States of America

TABLE OF CONTENTS

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

SECOND SERIES, VOL. 35, NO. 2
APRIL, 1934

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, \$5, a volume (four numbers) in advance. Single copies \$1.50. An arrangement has been made with the Mathematical Association of America in accordance with which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulae, and must be in final form with all references filled in.

Authors receive gratis 100 reprints of each article, postage prepaid.

Application has been made for entry as second class matter at the post office at Baltimore, Md.

Made in United States of America

Copies of the following memoirs can be obtained by addressing THE PRINCETON UNIVERSITY PRESS, Princeton, N. J.:

Fermat's last theorem and the origin and nature of the theory of algebraic numbers. By L. E. DICKSON. 27 pages. Price 35 cents.

An introduction to the theory of elliptic functions. By GOSTA MITTAG-LEFFLER. Authorized translation by EINAR HILLE. 81 pages. Price 90 cents.

Differential equations from the group standpoint. By L. E. DICKSON. 92 pages. Price $1.00.

On various conceptions of correlation. By F. M. WEIDA. 37 pages. Price 75 cents.

Contact transformations. By L. P. EISENHART. 40 pages. Price 80 cents.

A survey of the theory of small samples. By P. R. RIDER. 52 pages. Price $1.00.

Tauberian theorems. By NORBERT WIENER. 100 pages. Price $1.00.

Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension. By PAUL ALEXANDROFF. 86 pages. Price $1.00.

TABLE OF CONTENTS

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

S. LEFSCHETZ J. VON NEUMANN

WITH THE COÖPERATION OF THE

DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY

AND

THE SCHOOL OF MATHEMATICS OF THE INSTITUTE FOR ADVANCED STUDY

AND

SECOND SERIES, VOL. 35, NO. 3
JULY, 1934

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, $5, a volume (four numbers) in advance. Single copies $1.50. An arrangement has been made with the Mathematical Association of America in accordance with which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Md., or to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulae, and must be in final form with all references filled in.

Authors receive gratis 100 reprints of each article, postage prepaid.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in United States of America

TABLE OF CONTENTS

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

SECOND SERIES, VOL. 35, NO. 4
OCTOBER, 1934

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, \$5, a volume (four numbers) in advance. Single copies \$1.50. An arrangement has been made with the Mathematical Association of America in accordance with which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Md., or to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulae, and must be in final form with all references filled in.

Authors receive gratis 100 reprints of each article, postage prepaid.

Entered as second class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in United States of America

Copies of the following memoirs can be obtained by addressing The Princeton University Press, Princeton, N. J.:

Fermat's last theorem and the origin and nature of the theory of algebraic numbers. By L. E. Dickson. 27 pages. Price 35 cents.

An introduction to the theory of elliptic functions. By Gosta Mittag-Leffler. Authorized translation by Einar Hille. 81 pages. Price 90 cents.

Differential equations from the group standpoint. By L. E. Dickson. 92 pages. Price $1.00.

On various conceptions of correlation. By F. M. Weida. 37 pages. Price 75 cents.

Contact transformations. By L. P. Eisenhart. 40 pages. Price 80 cents.

A survey of the theory of small samples. By P. R. Rider. 52 pages. Price $1.00.

Tauberian theorems. By Norbert Wiener. 100 pages. Price $1.00.

Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension. By Paul Alexandroff. 86 pages. Price $1.00.

TABLE OF CONTENTS